PCT/GB99/04399

The Patent Office

INVESTOR IN PEOPLE

GB99
04399

REC'D 26 JAN 2000
WIPO PCT

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed R. Mahoney

Dated 17 January 2000

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Patents Form 1/77

Patents Act 1977
(Rule 16)

The Patent Office

29DEC98 E414681-4 D00060
P01/7700 0.00 - 9828715.4

THE PATENT OFFICE
24 DEC 1998
RECEIVED BY HAND

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference	SJW/BP5671631
2. Patent application number *(The Patent Office will fill in this part)*	**9828715.4**
3. Full name, address and postcode of the or of each applicant *(underline all surnames)*	UNIVERSITY COLLEGE LONDON GOWER STREET LONDON WC1E 6BT GB
Patents ADP number *(if you know it)*	
If the applicant is a corporate body, give the country/state of its incorporation	00798652002
4. Title of the invention	GLYCOSYLPHOSPHATIDYLINOSITOL SPECIFIC PHOSPHOLIPASE D PROTEINS AND USES THEREOF
5. Name of your agent *(if you have one)*	MEWBURN ELLIS
"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*	YORK HOUSE 23 KINGSWAY LONDON WC2B 6HP
Patents ADP number *(if you know it)*	109006

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application number	Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application	Number of earlier application	Date of filing *(day / month /year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:* — YES

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	0	
Description	33	✓ 85K
Claim*(s)*	0	
Abstract	0	
Drawing*(s)*	18	+18

10. If you are also filing any of the following, state

Priority documents	0
Translations of priority documents	0
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	0
Request for preliminary examination and search *(Patents Form 9/77)*	0
Request for substantive examination *(Patents Form 10/77)*	0
Any other documents *(Please specify)*	0

11. I/We request the grant of a patent on the basis of this application.

Signature Newburn Ellis

Date 23 December 1998

12. Name and daytime telephone number of person to contact in the United Kingdom — SIMON J. KIDDLE — 0117 926 6411

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed it it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

Glycosylphosphatidylinositol Specific Phospholipase D Proteins and Uses Thereof

Field of the Invention

The present invention relates to glycosylphosphatidylinositol specific phospholipase D (GPI-PLD) proteins and uses of these proteins, in particular in the treatment and diagnosis of diabetes and complications of diabetes such as insulin resistance.

Background of the Invention

Studies have shown that a number of cell surface proteins are attached to the cell membrane by covalent linkage to a glycosylphosphatidylinositol (GPI) anchor. It has been shown that the enzyme GPI-PLD cleaves the photodiester bond linking glycosylphosphatidylinositol to phosphatidic acid, thereby releasing anchored proteins.

GPI-PLD enzymes are abundantly present in human and bovine serum (5-10μg/ml in human serum). US Patent No: 5,418,147 (Huang et al) describes the purification of GPI-PLD from bovine liver, and the subsequent cloning of three GPI-PLD enzymes from bovine liver, human liver and human pancreas cDNA libraries. This patent reports the full length cDNA and amino acid sequences of the GPI-PLDs from human and bovine liver, and the partial cDNA and amino acid sequences of the human pancreatic form of the enzyme. Subsequently, the full length sequence of the pancreatic form of GPI-PLD was reported in Tsang et al (1992), and this enzyme has been found in cDNA libraries from breast, eye, spleen and tonsil. The three forms of the enzymes are highly homologous with the predicted mature protein sequences of bovine liver GPI-PLD sharing 82% sequence identity with the human liver enzyme and 81% sequence identity with the human pancreatic enzyme. The amino acid sequences of human liver and pancreatic forms of GPI-PLD were deposited at GenBank under accession numbers L11701 and L11702 and consist of 841 and 840 amino acids respectively. The human liver and pancreatic

forms of GPI-PLD share 94.6% sequence identity. The structure of GPI-PLDs is further discussed in Scallon et al, 1991.

However, despite cloning three forms of GPI-PLD, there is no suggestion in these references as to the *in vivo* role of the enzymes. Further, the only application of the enzymes suggested is in an expression system in which a heterologous protein is expressed in a host cell as a fusion with a GPI-signal peptide, leading to the heterologous protein becoming anchored to the cell membrane by a GPI anchor, where it can be cleaved off by coexpressed or added GPI-PLD.

GPI-PLD has also been isolated from human serum by Hoener et al (1992) and this form of the enzyme was found to be identical to the human pancreatic GPI-PLD apart from changes at 531 to 534 where VIGS is replaced by MLGT. This paper also showed that treatment of serum GPI-PLD with N-glycosidase F reduced the apparent molecular weight from 123 kD to 87 kD. Similarly, by Li et al (1994) showed GPI-PLD was cleaved by trypsin into 3 fragments (2 x 40 kD and 30 kD), and by Heller et al (1994) which showed that 33, 39 and 47kD species were produced, with only the N-terminal 39 kD fragment moiety showing enzyme activity after renaturation.

It has been proposed that one function of GPI-PLD enzyme is to produce inositolphosphoglycans (IPGs) by the cleavage of "free" GPIs in the plasma membrane in response to binding of a growth factor to its receptor (Rademacher et al, 1994). This role for GPI-PLD has been demonstrated in mast cells where IgE-dependent activation of these cells results in release of their granule contents, which include substances such as histamine, a mediator of the inflammatory response. In the presence of antigen, histamine is released; this release can be

mimicked by addition of IPGs and is blocked by addition of anti-GPI-PLD antibodies (Lin et al, 1991).

The role of GPI-PLD in cleaving GPI-anchored proteins, and especially inositolphosphoglycans (IPGs), is examined in Jones et al (1997). However, the authors reflect the uncertainty in the art regarding the mechanism of IPG generation, noting that "The definitive activated enzyme, being a GPI-PLC or a GPI-PLD, has yet to be unequivocally identified" and that "little attention has been payed to the role of GPI-PLD as the hydrolysing enzyme".

In summary, despite the cloning of GPI-PLD enzymes and investigation as to their biochemical properties, the role of the enzyme *in vivo* or any possible medical use remains unknown.

Summary of the Invention

Broadly, the present invention relates to GPI-PLD for medical use, and in particular to the use of GPI-PLD in the treatment of diabetes and complications of diabetes, optionally in combination with existing therapies, such as the administration of insulin. The GPI-PLD can be the forms of the enzyme disclosed in the prior art, or the GPI-PLDs disclosed for the first time here.

Insulin is a major anabolic hormone and has both mitogenic and metabolic effects. Whilst much effort has been directed towards study of the cascade of intracellular phosphorylation events initiated by the binding of insulin to its cell surface receptor, the signalling arm mediated by IPGs has been largely overlooked. The present invention is based on the realisation that GPI-PLDs are in fact the enzymes responsible for production of IPG second messengers following the binding of insulin to its receptor. The IPGs then interact with other cellular enzymes

instigating some of the metabolic effects of the hormone. In particular, diabetic complications such as insulin resistance may be caused by deficiencies in GPI-PLD. Pancreatic islet cells produce and secrete GPI-PLD, which is transported in blood complexed with apolipoprotein A1, and may therefore represent the major source of circulating enzyme.

Insulin resistance is seen in both the early stages of type I (IDDM) and type II diabetes mellitus (NIDDM). If GPI-PLD levels are depleted by the destruction of pancreatic β-cells, as is seen in streptozotocin-treated rats, then this could be an important factor in the development of insulin resistance. This in turn suggests the treatment of such patients with GPI-PLD, optionally in combination with other diabetes therapies.

Accordingly, in a first aspect, the present invention provides GPI-PLD for use in a method of medical treatment.

In a further aspect, the present invention provides the use of GPI-PLD for the preparation of a medicament for the treatment of diabetes, and in particular insulin dependent forms of diabetes.

In a further aspect, the present invention provides the use of GPI-PLD for the preparation of a medicament for the treatment of complications of diabetes, and in particular the treatment of insulin resistance.

In a further aspect, the present invention provides a method of treating a patient having diabetes or complications arising from diabetes, the method comprising administering to the patient a therapeutically effective amount of GPI-PLD.

In all of the above aspects, GPI-PLD can be administered alone or in conjunction with other treatments for diabetes or diabetic complications, either sequentially or simultaneously.

In a further aspect, the present invention provides a kit comprising a composition including GPI-PLD and a second composition for the treatment of diabetes.

In a further aspect, the present invention provides an expression vector comprising nucleic acid encoding GPI-PLD for use in a method of gene therapy, e.g. in the treatment of patients unable to produce sufficient GPI-PLD. The GPI-PLD encoding nucleic acid can be a sequence shown in figures 4 to 6 or one of the known nucleic acid sequences.

In a further aspect, the present invention provides a cell line for transplantation into a patient, wherein the cell line is transformed with nucleic acid encoding GPI-PLD, and is capable of expressing and secreting GPI-PLD. In one embodiment, the cell line is encapsulated, e.g. in a biocompatible polymer, so that the GPI-PLD produced by the cell line can be secreted into the patient, while preventing rejection by the immune system of the host. Methods for encapsulating cells in biocompatible polymers are described in WO93/16687 and WO96/31199.

In a further aspect, the present invention provides a pharmaceutical composition comprising a nucleic acid molecule encoding a GPI-PLD protein.

In a further aspect, the present invention provides a pharmaceutical composition comprising a GPI-PLD protein.

In a further aspect, the present invention provides the use of GPI-PLD levels in the diagnosis of diabetes or

diabetic complications. Thus, the present invention provides a method of diagnosing diabetes or diabetic complications, the method comprising determining the amount of GPI-PLD in a biological sample from a patient. This determination can help in the diagnosis or prognosis of the patient, allowing the treatment of the patient to be tailored accordingly to the patient's individual needs.

In one embodiment, the method of diagnosing diabetes or diabetic complications comprises the steps of:

(a) contacting a biological sample obtained from the patient with a solid support having immobilised thereon a binding agent having binding sites specific for GPI-PLD;

(b) contacting the solid support with one or more labelled developing agents capable of binding to unoccupied binding sites, bound GPI-PLD or occupied binding sites; and,

(c) detecting the label of the developing agents specifically binding in step (b) to obtain a value representative of the amount of GPI-PLD in the sample.

Alternatively or additionally, the method can assess GPI-PLD levels by measuring one of its biological activities, which are discussed further below.

The present invention also relates to novel GPI-PLD proteins and nucleic acid molecules, and in particular to forms of the protein having sequence differences compared to the known human liver and pancreatic forms reported in the prior art.

In a further aspect, the present invention provides a substance which is an isolated polypeptide comprising a polypeptide having the amino acid sequence set out in figure 3.

In a further aspect, the present invention provides isolated nucleic acid molecules encoding any one of the above polypeptides. Examples of such nucleic acid sequences are the nucleic acid sequences set out in figures 4 to 6. The present invention also include nucleic molecules having greater than 90% sequence identity with the nucleic acid sequences shown in these figures.

In further aspects, the present invention provides an expression vector comprising the above GPI-PLD proteins, nucleic acid operably linked to control sequences to direct its expression, and host cells transformed with the vectors. The present invention also includes a method of producing the above GPI-PLD proteins comprising culturing the host cells and isolating the GPI-PLD thus produced.

These and other aspects of the present invention are described in more detail below.

By way of example, embodiments of the present invention will now be described in more detail with reference to the accompanying figures.

Brief Description of the Figures

Figure 1 shows an alignment of the deduced amino acid sequences of GPI-PLD encoded by cDNA clone A1 and the bovine and human liver GPI-PLD sequences disclosed in US Patent No: 5,418,147 (Huang et al).

Figure 2 shows the nucleic acid sequence from cDNA clone A1 aligned with the pancreatic forms of GPI-PLD disclosed in US Patent No: 5,418,147 (Huang et al) (partial sequence) and the corresponding full length nucleic acid sequence deposited at GenBank.

Figure 3 shows the amino acid sequences of the GPI-PLDs in clones a1, b2 and d3, and consist of 840, 795 and 510 amino acids respectively.

Figure 4 shows the nucleic acid sequence of cDNA clone a1 encoding GPI-PLD, consisting of 2832 bp.

Figure 5 shows the nucleic acid sequence of cDNA clone b2 encoding GPI-PLD, consisting of 2472 bp.

Figure 6 shows the nucleic acid sequence of cDNA clone d3 encoding GPI-PLD, consisting of 1942 bp.

Figure 7 shows an alignment of the deduced amino acid sequences of GPI-PLDs encoded by cDNA clones a1, b2 and d3 with the pancreatic form of the enzyme deposited at GenBank under accession number 11702.

Figure 8 shows an alignment of the nucleic acid sequences from cDNA clones a1, b2 and d3 with the cDNA sequence encoding the human pancreatic form of GPI-PLD deposited at GenBank under accession number 11702.

Detailed Description

GPI-PLD Proteins

The term "GPI-PLD biological activity" is herein defined as the enzymatic activity of GPI-PLD in cleaving the photodiester bond linking glycosylphosphatidylinositol to phosphatidic acid, e.g. releasing a GPI-anchored protein. As noted in Heller et al (1994), this activity has been localised to the N-terminal 39 kD portion of full length GPI-PLD.

The medical uses of GPI-PLD described herein can use the novel GPI-PLD variants or the forms of the enzyme disclosed in the prior art. In either event, the skilled person can use the techniques described herein and others

well known in the art to produce large amounts of these proteins, or fragments or active portions thereof, for use as pharmaceuticals, in the developments of drugs and for further study into its properties and role *in vivo*.

In a further aspect of the present invention provides a polypeptide having the amino acid sequence shown in figure 3, which may be in isolated and/or purified form, free or substantially free of material with which it is naturally associated. In one embodiment, the clone a1 has an amino acid sequence consisting of 840 amino acids, a 23 amino acid signal peptide and a 817 amino acid mature protein.

GPI-PLD proteins which are amino acid sequence variants, alleles or derivatives can also be used in the present invention. A polypeptide which is a variant, allele or derivative may have an amino acid sequence which differs from that given in figures 1 or 3 by one or more of addition, substitution, deletion and insertion of one or more amino acids. Preferred polypeptides have GPI-PLD enzymatic function as defined above.

A GPI-PLD protein which is an amino acid sequence variant, allele or derivative of an amino acid sequence shown in figures 1 or 3 may comprise an amino acid sequence which shares greater than about 70%, greater than about 80%, greater than about 90%, greater than about 95%, greater than about 97%, greater than about 98% or greater than about 99% sequence identity with an amino acid sequence shown in figures 1 or 3. Sequence comparison and identity calculations were carried out using the Cluster program (Thompson et al, 1994), using the following parameters (Pairwise Alignment Parameters: Weight Matrix: pam series; Gap Open Penalty: 10.00; Gap Extension Penalty: 0.10). Alternatively, the GCG program could be used which is available from Genetics Computer

Group, Oxford Molecular Group, Madison, Wisconsin, USA, Version 9.1. Particular amino acid sequence variants may differ from those shown in figures 1 and 3 by insertion, addition, substitution or deletion of 1 amino acid, 2, 3, 4, 5-10, 10-20 20-30, 30-50, 50-100, 100-150, or more than 150 amino acids.

The present invention also includes the use of active portions, fragments and derivatives of the GPI-PLD proteins.

An "active portion" of GPI-PLD protein is a polypeptide which is less than said full length GPI-PLD protein, but which retains at least one its essential biological activity, e.g. the enzyme activity mentioned above. For instance, portions of GPI-PLD protein can act as sequestrators or competitive antagonists by interacting with other proteins.

A "fragment" of the GPI-PLD protein means a stretch of amino acid residues of at least about 5 to 7 contiguous amino acids, often at least about 7 to 9 contiguous amino acids, typically at least about 9 to 13 contiguous amino acids and, most preferably, at least about 20 to 30 or more contiguous amino acids.

A "derivative" of the GPI-PLD protein, or a fragment thereof, means a polypeptide modified by varying the amino acid sequence of the GPI-PLD protein, e.g. by manipulation of the nucleic acid encoding the protein or by altering the protein itself. Such derivatives of the natural amino acid sequence may involve insertion, addition, deletion or substitution of one, two, three, five or more amino acids, without fundamentally altering a biological activity of the wild type GPI-PLD protein.

A polypeptide according to the present invention may be

isolated and/or purified (e.g. using an antibody) for instance after production by expression from encoding nucleic acid (for which see below). Polypeptides according to the present invention may also be generated wholly or partly by chemical synthesis. The isolated and/or purified polypeptide may be used in formulation of a composition, which may include at least one additional component, for example a pharmaceutical composition including a pharmaceutically acceptable excipient, vehicle or carrier. A composition including a polypeptide according to the invention may be used in prophylactic and/or therapeutic treatment as discussed below.

The GPI-PLD polypeptides can also be linked to a coupling partner, e.g. an effector molecule, a label, a drug, a toxin and/or a carrier or transport molecule. Techniques for coupling the peptides of the invention to both peptidyl and non-peptidyl coupling partners are well known in the art. In one embodiment, the carrier molecule is a 16 aa peptide sequence derived from the homeodomain of *Antennapedia* (e.g. as sold under the name "Penetratin"), which can be coupled to a peptide via a terminal Cys residue. The "Penetratin" molecule and its properties are described in WO91/18981.

Pharmaceutical Compositions

As mentioned above, GPI-PLD proteins can used for treating diabetes and the complications of diabetes (e.g. insulin resistance), optionally in conjunction with other treatments for these disorders. Thus, the GPI-PLD proteins can be formulated in pharmaceutical compositions, which may comprise, in addition to one of the above substances, a pharmaceutically acceptable excipient, carrier, buffer, stabiliser or other materials well known to those skilled in the art. Such materials should be non-toxic and should not interfere with the

efficacy of the active ingredient. The precise nature of the carrier or other material may depend on the route of administration, e.g. oral, intravenous, cutaneous or subcutaneous, nasal, intramuscular, intraperitoneal routes.

Pharmaceutical compositions for oral administration may be in tablet, capsule, powder or liquid form. A tablet may include a solid carrier such as gelatin or an adjuvant. Liquid pharmaceutical compositions generally include a liquid carrier such as water, petroleum, animal or vegetable oils, mineral oil or synthetic oil. Physiological saline solution, dextrose or other saccharide solution or glycols such as ethylene glycol, propylene glycol or polyethylene glycol may be included.

For intravenous, cutaneous or subcutaneous injection, or injection at the site of affliction, the active ingredient will be in the form of a parenterally acceptable aqueous solution which is pyrogen-free and has suitable pH, isotonicity and stability. Those of relevant skill in the art are well able to prepare suitable solutions using, for example, isotonic vehicles such as sodium chloride injection, Ringer's injection, lactated Ringer's injection. Preservatives, stabilisers, buffers, antioxidants and/or other additives may be included as required.

Whether it is a polypeptide, peptide, nucleic acid molecule, small molecule or other pharmaceutically useful compound of the invention that is to be given to an individual, administration is preferably in a "prophylactically effective amount" or a "therapeutically effective amount" (as the case may be, although prophylaxis may be considered therapy), this being sufficient to show benefit to the individual. The actual amount administered, and rate and time-course of

administration, will depend on the nature and severity of what is being treated. Prescription of treatment, e.g. decisions on dosage etc, is within the responsibility of general practitioners and other medical doctors, and typically takes account of the disorder to be treated, the condition of the individual patient, the site of delivery, the method of administration and other factors known to practitioners. Examples of the techniques and protocols mentioned above can be found in Remington's Pharmaceutical Sciences, 16th edition, Oslo, A. (ed), 1980.

GPI-PLD proteins can be administered alone or in combination with other treatments for diabetes or diabetic complications, either simultaneously or sequentially. Examples of known diabetes treatments include (a) insulin, which is typically delivered by injection, (b) oral insulin compositions, (c) glucose sparing or insulin enhancing drugs, (d) α-glucosidase inhibitors to reduce carbohydrate absorption (precose and miglitol), and (e) drugs used to treat patients with insulin sensitivity, e.g. thiazolidinediones, such as Rezulin, rosiglitazone, piogliazone and tyrosine phosphatase inhibitors.

In further embodiments, the GPI-PLD can be administered with P and/or A-type IPGs, and/or antagonists of these substances. Methods for obtaining A-type and P-type IPGs and their antagonists are set out in Caro et al, 1997, and in WO98/11116 and WO98/11117.

The role of P and A-type IPGs and their use in the diagnosis and treatment of diabetes is disclosed in WO98/11435. In summary, this application discloses that in some forms of diabetes the ratio of P:A type IPGs is imbalanced and can be corrected by administering a medicament comprising the appropriate ratio of P or A-

type IPGs or antagonist thereof. In particular, WO98/11435 describes the treatment of obese type II diabetes (NIDDM) patients with a P-type IPG or with an A-type IPG antagonist, such as antibodies which bind specifically to the A-type IPG, and the treatment of IDDM or lean type II diabetes (NIDDM) (body mass index < 27) with a mixture of A and P-type IPGs, typically in a P:A ratio of about 6:1 for males and 4:1 for females.

The compositions of the invention can be used in the treatment of diabetes, in particular insulin dependent forms of diabetes (type I and type II diabetes). They can also be used in the treatment of the complications of diabetes and in particular forms of insulin resistance such as insulin resistance in type I or type II diabetes and brittle diabetes. The compositions may also be used to treat other conditions mediated by insulin, and in particular insulin underproduction such as neutrotrophic disorders or polycystic ovary disease.

GPI-PLD nucleic acid

"GPI-PLD nucleic acid" includes a nucleic acid molecule which has a nucleotide sequence encoding a polypeptide which includes the amino acid sequence shown in figures 4 to 6, and in some embodiments of the invention extends to the known human liver and pancreatic forms of GPI-PLD (L11701 and L11702). These forms of GPI-PLD have been mapped to human chromosome 6 and are contained in the 4 centimorgan region of D6S1660-D6S1558 at positions 95.95 and 99.71 (NCBI GeneMap'98). This corresponds to the cytogenetic region of 6p21.3. This region also contains the IDDM1 and HLA loci (although the HLA genes map to the adjacent D6S1558-D6S1616 interval). The mouse GPI-PLD gene has also been mapped to chromosome 13, near the *fim 1* locus, which is found in humans on chromosome 6.

The GPI-PLD coding sequence may be that shown in figures

2, 4 to 6 or 8, a complementary nucleic acid sequence, or it may be a mutant, variant, derivative or allele of these sequences. The sequence may differ from that shown by a change which is one or more of addition, insertion, deletion and substitution of one or more nucleotides of the sequence shown. Changes to a nucleotide sequence may result in an amino acid change at the protein level, or not, as determined by the genetic code.

The encoded polypeptide may comprise an amino acid sequence which differs by one or more amino acid residues from the amino acid sequence shown in the figures. Nucleic acid encoding a polypeptide which is an amino acid sequence mutant, variant, derivative or allele of the sequence shown in figures 1, 3 or 7 is further provided by the present invention. Such polypeptides are discussed below. Nucleic acid encoding such a polypeptide may show greater than about 70% identity, greater than about 80% identity, greater than about 90% identity, greater than about 95% identity, greater than about 98% identity, or greater than about 99% identity with a sequence shown in the figures.

The present invention also includes fragments of the GPI-PLD nucleic acid sequences described herein, the fragments preferably being at least 12, 15, 30, 45, 60, or 120 nucleotides in length.

Generally, nucleic acid according to the present invention is provided as an isolate, in isolated and/or purified form, or free or substantially free of material with which it is naturally associated, such as free or substantially free of nucleic acid flanking the gene in the human genome, except possibly one or more regulatory sequence(s) for expression. Nucleic acid may be wholly or partially synthetic and may include genomic DNA, cDNA or RNA. Where nucleic acid according to the invention

includes RNA, reference to the sequence shown should be construed as reference to the RNA equivalent, with U substituted for T.

Nucleic acid sequences encoding all or part of the GPI-PLD gene and/or its regulatory elements can be readily prepared by the skilled person using the information and references contained herein and techniques known in the art (for example, see Sambrook, Fritsch and Maniatis, "Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989, and Ausubel et al, Short Protocols in Molecular Biology, John Wiley and Sons, 1992). These techniques include (i) the use of the polymerase chain reaction (PCR) to amplify samples of such nucleic acid, e.g. from genomic sources, (ii) chemical synthesis, or (iii) amplification in *E. coli*. Modifications to the GPI-PLD sequences can be made, e.g. using site directed mutagenesis, to provide expression of modified GPI-PLD protein or to take account of codon preference in the host cells used to express the nucleic acid.

In order to obtain expression of the GPI-PLD nucleic acid sequences, the sequences can be incorporated in a vector having control sequences operably linked to the GPI-PLD nucleic acid to control its expression. The use of expression systems has reached an advanced degree of sophistication. The vectors may include other sequences such as promoters or enhancers to drive the expression of the inserted nucleic acid, nucleic acid sequences so that the GPI-PLD protein is produced as a fusion and/or nucleic acid encoding secretion signals so that the polypeptide produced in the host cell is secreted from the cell. GPI-PLD protein can then be obtained by transforming the vectors into host cells in which the vector is functional, culturing the host cells so that the GPI-PLD protein is produced and recovering the GPI-

PLD protein from the host cells or the surrounding medium. Prokaryotic and eukaryotic cells are used for this purpose in the art, including strains of *E. coli*, yeast, and eukaryotic cells such as COS or CHO cells. The choice of host cell can be used to control the properties of the GPI-PLD protein expressed in those cells, e.g. controlling where the polypeptide is deposited in the host cells or affecting properties such as its glycosylation and phosphorylation.

PCR techniques for the amplification of nucleic acid are described in US Patent No: 4,683,195. In general, such techniques require that sequence information from the ends of the target sequence is known to allow suitable forward and reverse oligonucleotide primers to be designed to be identical or similar to the polynucleotide sequence that is the target for the amplification. PCR comprises steps of denaturation of template nucleic acid (if double-stranded), annealing of primer to target, and polymerisation. The nucleic acid probed or used as template in the amplification reaction may be genomic DNA, cDNA or RNA. PCR can be used to amplify specific sequences from genomic DNA, specific RNA sequences and cDNA transcribed from mRNA, bacteriophage or plasmid sequences. The GPI-PLD protein nucleic acid sequences provided herein readily allow the skilled person to design PCR primers. References for the general use of PCR techniques include Mullis et al, Cold Spring Harbor Symp. Quant. Biol., 51:263, 1987; Ehrlich (ed), PCR Technology, Stockton Press, NY, 1989; Ehrlich et al, Science, 252:1643-1650, 1991; "PCR protocols; A Guide to Methods and Applications", Eds. Innis et al, Academic Press, New York, 1990.

Nucleic acid according to the present invention is obtainable using one or more oligonucleotide probes or primers designed to hybridize with one or more fragments

of the nucleic acid sequence shown in the figures, particularly fragments of relatively rare sequence, based on codon usage or statistical analysis. A primer designed to hybridize with a fragment of the nucleic acid sequence shown in the above figures may be used in conjunction with one or more oligonucleotides designed to hybridize to a sequence in a cloning vector within which target nucleic acid has been cloned, or in so-called "RACE" (rapid amplification of cDNA ends) in which cDNA's in a library are ligated to an oligonucleotide linker and PCR is performed using a primer which hybridizes with a GPI-PLD nucleic acid sequence shown in figures and a primer which hybridizes to the oligonucleotide linker.

Such oligonucleotide probes or primers, as well as the full-length sequence (and mutants, alleles, variants and derivatives) are also useful in screening a test sample containing nucleic acid for the presence of alleles, mutants and variants, especially those that lead to the production of inactive forms of GPI-PLD protein protein, the probes hybridizing with a target sequence from a sample obtained from the individual being tested. The conditions of the hybridization can be controlled to minimise non-specific binding, and preferably stringent to moderately stringent hybridization conditions are preferred. The skilled person is readily able to design such probes, label them and devise suitable conditions for the hybridization reactions, assisted by textbooks such as Sambrook et al (1989) and Ausubel et al (1992).

Examples of "stringent conditions" are those which: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulphate at 50°C; (2) employ during hybridisation a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% BSA/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium

phosphate buffer at pH 6.5 with 750mM sodium chloride, 75mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50μg/ml), 0.1% SDS, and 10% dextran sulphate at 42°C, with washes at 42°C in 0.2 x SSC and 50% formamide at 55°C, followed by high stringency wash consisting of 0.1 x SSC containing EDTA at 55°C. These hybridisation conditions may be used in the context of defining nucleic acid sequences which hybridize with GPI-PLD nucleic acid sequences.

<u>**Uses of GPI-PLD Nucleic Acid**</u>

The GPI-PLD nucleic acid sequences can be used in the preparation of cell lines capable of expressing GPI-PLD and in gene therapy techniques.

Thus, the present invention provides a cell line for transplantation into a patient, the cell line being transformed with nucleic acid encoding GPI-PLD, and being capable of expressing and secreting GPI-PLD. In one embodiment, the cell lines are encapsulated, e.g. in a biocompatible polymer, so that the GPI-PLD produced by the cell line can be secreted into the patient, while preventing rejection by the immune system of the host. Methods for encapsulating cells in biocompatible polymers are described in WO93/16687 and WO96/31199.

As a further alternative, the nucleic acid encoded the GPI-PLD protein could be used in a method of gene therapy, to treat a patient who is unable to synthesize the active polypeptide or unable to synthesize it at the normal level, thereby providing the effect provided by wild-type GPI-PLD protein and suppressing the occurrence of diabetes in the target cells.

Vectors such as viral vectors have been used in the prior art to introduce genes into a wide variety of different target cells. Typically, the vectors are exposed to the target cells so that transfection can take place in a sufficient proportion of the cells to provide a useful therapeutic or prophylactic effect from the expression of the desired polypeptide. The transfected nucleic acid may be permanently incorporated into the genome of each of the targeted tumour cells, providing long lasting effect, or alternatively the treatment may have to be repeated periodically.

A variety of vectors, both viral vectors and plasmid vectors, are known in the art, see US Patent No: 5,252,479 and WO93/07282. In particular, a number of viruses have been used as gene transfer vectors, including papovaviruses, such as SV40, vaccinia virus, herpesviruses, including HSV and EBV, and retroviruses. Many gene therapy protocols in the prior art have used disabled murine retroviruses.

As an alternative to the use of viral vectors other known methods of introducing nucleic acid into cells includes electroporation, calcium phosphate co-precipitation, mechanical techniques such as microinjection, transfer mediated by liposomes and direct DNA uptake and receptor-mediated DNA transfer.

As mentioned above, the aim of gene therapy using nucleic acid encoding the GPI-PLD protein, or an active portion thereof, is to increase the amount of the expression product of the nucleic acid in cells in which the level of the wild-type GPI-PLD protein is absent or present only at reduced levels. Target cells for gene therapy include insulin secreting β-cells or any neuron derived cells. Cell engineering can be used to provide the overexpression or repression of GPI-PLD protein in

transfected cell lines which can then be subsequently transplanted to humans. Gene therapy can be employed using a promoter to drive GPI-PLD protein expression in a tissue specific manner (i.e. an insulin promoter linked to GPI-PLD cDNA will overexpress GPI-PLD protein in β-cells and transiently in the brain). If defective function of GPI-PLD protein is involved in neurological disease, GPI-PLD protein can be overexpressed in transformed cell lines for transplantation.

Gene transfer techniques which selectively target the GPI-PLD nucleic acid to target tissues are preferred. Examples of this included receptor-mediated gene transfer, in which the nucleic acid is linked to a protein ligand via polylysine, with the ligand being specific for a receptor present on the surface of the target cells.

<u>Diagnostic Methods</u>

Methods for determining the concentration of analytes in biological samples from individuals are well known in the art and can be employed in the context of the present invention to determine the presence or amount of GPI-PLD in a biological sample from a patient. This in turn can allow a physician to determine whether a patient suffers from diabetes or diabetic complications, and so optimise the treatment of it.

Broadly, the methods divide into those screening for the presence of GPI-PLD protein nucleic acid sequences and those that rely on detecting the presence or absence of the GPI-PLD protein polypeptide. The methods make use of biological samples from individuals that are suspected of contain the nucleic acid sequences or polypeptide.

These diagnostic methods can employ biological samples such as blood, serum, tissue samples or urine. In view

of the fact that the activity of GPI-PLD is thought to be due to the level of the enzyme circulating in serum, the use of serum or blood samples is preferred.

The assay methods for determining the amount or concentration of GPI-PLD protein typically either employ binding agents having binding sites capable of specifically binding to GPI-PLD in preference to other molecules or measure a characteristic biological activity of GPI-PLD. Examples of binding agents include antibodies, receptors and other molecules capable of specifically binding the enzyme. Conveniently, the binding agent(s) are immobilised on solid support, e.g. at defined locations, to make them easy to manipulate during the assay.

The sample is generally contacted with the binding agent(s) under appropriate conditions so that GPI-PLD present in the sample can bind to the binding agent(s). The fractional occupancy of the binding sites of the binding agent(s) can then be determined using a developing agent or agents. Typically, the developing agents are labelled (e.g. with radioactive, fluorescent or enzyme labels) so that they can be detected using techniques well known in the art. Thus, radioactive labels can be detected using a scintillation counter or other radiation counting device, fluorescent labels using a laser and confocal microscope, and enzyme labels by the action of an enzyme label on a substrate, typically to produce a colour change. The developing agent(s) can be used in a competitive method in which the developing agent competes with the analyte for occupied binding sites of the binding agent, or non-competitive method, in which the labelled developing agent binds analyte bound by the binding agent or to occupied binding sites. Both methods provide an indication of the number of the binding sites occupied by the analyte, and hence the

concentration of the analyte in the sample, e.g. by comparison with standards obtained using samples containing known concentrations of the analyte.

Experimental

The present invention is based on the realisation that GPI-PLD is responsible for the production of IPG second messengers following binding of insulin to its receptor. The IPGs then interact with other cellular enzymes instigating some of the metabolic effects of the hormone. In view of this, insulin resistance may be caused by deficiencies in GPI-PLD; it has shown that pancreatic islet cells produce and secrete GPI-PLD, which is transported in blood complexed with apolipoprotein A1, and may therefore represent the major source of circulating enzyme. If this is indeed the case then the insulin resistance seen in early type I diabetes mellitus (IDDM) may result from decreased circulating GPI-PLD levels. This may have direct therapeutic relevance in that co-infusion of insulin with GPI-PLD may in fact be a far more effective therapy for diabetic patients than insulin.

Screening of human liver cDNA library

A human liver cDNA library (Gibco BRL, cat # 10422-012, lot # HF4703) was screened for GPI-PLD, resulting in the isolation of 3 cDNA clones. The nucleic acid sequences of the clones are shown in figures 4 to 6, with the deduced amino acid sequences shown in figure 3.

Clone a1 represents the full length cDNA. There are only two differences within the coding region of this sequence when compared to that of the human GPI-PLD pancreatic form described in the GenBank database (accession number L11702). These are a *g* to *a* conversion at positions 88 (L11702), 199 (a1) and a *t* to *g* conversion at positions 797 (L11702), 908(a1). Interestingly this latter this

latter conversion creates a unique *HindIII* restriction site in the a1 clone. Both conversions result in amino acid differences, the first changes amino acid 30 from a valine in L11702 to an isoleucine in a1, and the second changes amino acid 266 from an isoleucine in L11702 to a serine in a1. Clone a1 also differs from L11702 in that it contains 5′ untranslated region (UTR) and only shares the first 168 bases of the 3′ UTR before terminating in a poly-A tail.

Clone b2 lacks the exon of GPI-PLD, which begins at position 2469 in the a1 nucleotide sequence. However, the sequence from here to the end of b2 (2444-2473) does not contain a stop codon. It is therefore not clear whether b2 represents a cDNA with a different final exon or is the produce of aberrant processing.

Clone d3 shared the coding 3′ UTR sequence of the a1 clone from a1 position 1119 onwards, however the initial 1008 base pairs of coding sequence are absent from this clone. Clone d3 contains a methionine initiation codon in frame to the coding sequence at position 202 and a unique 5′ UTR. Translation of d3 from this codon would result in a unique sequence of 6 amino acids (1-6). Clone d3 therefore appears to represent a true transcript, in that it contains initiation and stop codons and both 5′ and 3′ UTRs. The predicted protein product of this transcript would apparently lack the catalytic domain, which has been localised to the N-terminus of the GPI-PLD enzyme (amino acids 1-375), however the 3 EF hand-like domains would still be present.

Huang et al and Tsang et al (1992) reported that two variants or isoenzymes of GPI-PLD exist, the so-called liver and pancreatic forms (accession numbers L11701 and 11702). Other workers have detected L11702 cDNAs in

human breast, eye, spleen, tonsil, and pancreas, as well as in liver. However, we failed to detect the liver form of GPI-PLD in the liver or in any other tissues.

Gene mapping and localisation

The chromosomal gene isolated in the experiments above is about 20-30 kb in length. The gene was also isolated on a PAC and mapped by fluorescence-in situ hybridisation (FISH) to 6p21.3, agreeing with recent radiation hybrid maps as seen on GeneMap'98; NCBI). The IDDM1 susceptibility gene also maps to 6p21.3, although recent evidence suggests that at least two closely-linked loci for IDDM1 are in the MHC region. The MHC locus itself seems to map to a region adjoining the GPI-PLD locus rather than within the same microsatellite band, so the significance of the proximity of the GPI-PLD and IDDM1 loci is unclear.

Northern blots of the mRNA species found in liver have shown two presumed splice variants as well as the full-length transcript. One has a deletion of about 160 amino acids from the mature 817 amino acid protein. The other seems to be a C-terminal deletion, which may therefore be non-functional if other authors are correct in finding that the C-terminus is necessary for enzyme activity.

The predominant GPI-PLD species detected after tissue extraction by antibodies (Western blots) has apparent molecular weight of about 47 kD, which agrees with other authors that full-length GPI-PLD is taken up from the plasma and processed to smaller active species.

GPI-PLD obtained from serum by cells is required for second messenger signalling

The principle goal of these experiments was to determine the role of glycosylphosphatidylinositol phospholipase D (GPI-PLD) in a type one hypersensitivity reaction. This

reaction involved the cross-linking of IgE receptors on the mast cell surface, leading to the release of allergic mediators.

Such an allergic reaction has been experimentally reproduced in our laboratory, using a rat basophilic leukaemia cell line, RBL-2H3. These cells naturally have unoccupied IgE receptors (FcεR1, or high-affinity receptors), allowing them to be passively sensitised with an IgE isotype of choice.

RBL-2H3 cell culture was maintained in Eagles minimum essential medium, containing 10% Foetal Bovine Serum (FBS) (heat activated), 100 U/ml Penicillin, 100 μg/ml Streptomycin and 2 mM L-glutamine.

Previous research indicates that RBL-2H3 cells derive their GPI-PLD from the culture serum (data not shown). Therefore, it follows that inactivation of this external source of GPI-PLD would deprive the cells of any further enzyme.

Inactivation of GPI-PLD activity in foetal bovine serum was achieved according to the method of Kung et al (Biochimica et Biophysica Acta, 1357:329-338, 1997). Briefly, FCS was adjusted to pH 11 using concentrated hydrochloric acid, and incubated for 1 hour at 37°C using. After this time, the pH was adjusted to 7.4, and GPI-PLD activity was determined using an enzymatic assay (Davitz et al, J. Biol. Chem., 264:13760-13764, 1989). Results indicated that this alkaline incubation severely depleted GPI-PLD activity (data not shown).

To determine the effect of culture of RBL-2H3 cells in GPI-PLD inactive serum, the supplemented MEM was replaced with MEM in which the FBS had been inactivated. Although the cell appearance was not dramatically altered by the

altered culture conditions, determination of GPI-PLD activity showed a dramatic reduction in activity.

GPI-PLD activity in cells cultured with GPI-PLD active/inactive FBS:

Active = 0.66 units GPI-PLD activity/mg of protein.

Inactive = 0.11 units GPI-PLD activity/mg of protein.

The effect of a reduced GPI-PLD activity on the cell's ability to respond to IgE cross-linking was determined as follows:

RBL-2H3 cells were grown to confluence, after which time the adherent cells were removed from the culture flask using a cell scraper. The cell density was determined, using a haemocytometer, and adjusted to 2×10^5 per ml. The cells were seeded at 1 ml per well in a 24 well culture plate and cultured for overnight at 37°C in a humidified 5% CO_2 incubator.

The overnight culture media was aspirated and replaced with fresh media containing Rat IgE anti-DNP 3µg/ml. After a 2 hour incubation period, the media was aspirated, and the cells were washed twice, with HEPES Buffered Saline. Cross-linking was achieved by the addition of 200 µl of DNP-Albumin at 100 ng/ml, and incubation for 2 hours. Mediator release was determined using a colorimetric assay to detect the presence of β-hexosaminidase and compared with the total cell β-hexosaminidase content (as determined by incubation with 200 µl 5% Triton X-100 detergent). (Yasuda et al, Int. Imunol., 7:251-258, 1995). As shown in the table below, the responsiveness to cross-linking was significantly reduced in those cells that were cultured in GPI-PLD inactive media.

Percentage release in IgE linking activity assay (compared with total)

Active GPI-PLD culture = 48.79%

Inactive GPI-PLD culture = 5.07%

Phosphorylation of GPI-PLD

The phosphorylation state of the GPI-PLD enzymes can be determined using MALDI-TOF mass spectrometry as described by Yip & Hutchins (1992). Spectrums of tryptic digests of the four proteins can be compared before and after treatment with calf intestinal alkaline phosphatase. The specific kinases responsible for phosphorylation of GPI-PLD can then be determined by incubation of the GPI-PLD tryptic fragments with ATP in the presence of various kinases. Motif analysis of the amino acid sequence of human GPI-PLD using the HGMP "motif" package has revealed the presence of numerous potential phosphorylation sites for two enzymes: protein kinase C and protein kinase ck2 (formerly known as casine kinase II). These enzymes may therefore be involved in the activation of GPI-PLD. Intriguingly the activity of protein kinase ck2 has been shown to be modulated by IPGs (Alemany et al, 1990) and there is also indirect evidence suggesting that IPGs may act through protein kinase C, thus suggesting the possibility of feedback loops regulating the production of IPGs.

GPI-PLD as a metal ion transferase

Two families of IPGs exist. IPGs of the P-type stimulate incorporation of glucose into glycogen whereas the A-type IPGs stimulate incorporation of glucose into lipid. Metal ion analysis has shown that the P-type IPGs contain manganese and the A-type zinc. It is known that the serum form of GPI-PLD contains approximately 10 atoms of

zinc per molecule. Investigation can therefore show whether the different isoforms of human GPI-PLD produce IPGs with differing metal ion content.

This experiment can be performed in two ways. Firstly purified A-type and P-type IPGs can be extracted from rat liver (Caro et al, 1997) and their metal ions removed using dithiazone in chloroform. The IPGs can be incubated in the presence of radiosotopes of zinc ($^{65}Zn^{2+}$) and manganese ($^{52}Mn^{2+}$) respectively. The radiolabelled IPGs can then be added to the different isoforms of purified GPI-PLD (as determined in the above experiments) in the absence of GPI substrate thus driving the reaction from product (IPG) to substrate (GPI). It can then be determined whether or not the GPI-PLD protein have incorporated radioactive metal ions from the IPGs. The reverse situation will also be examined, whereby the metal ions of GPI-PLD isoforms are replaced by the respective radioisotopes. GPI-PLD can then be incubated with GPIs extracted from membrane preparations and the resulting IPG products analysed for incorporation of radioisotope. These experiments will thus determine whether or not GPI-PLD is responsible for the transfer of divalent cations (Mn^{2+} or Zn^{2+}) to its IPG products.

The significance of co-secretion of GPI-PLD and insulin from islet cells

It has been reported that the pancreatic islet B cell-derived mouse cell line BTC3 co-secretes insulin and GPI-PLD. Whilst the kinetics of secretion were somewhat different, both proteins were co-localised in secretory granules and their release was stimulated by glucose and a variety of secretagogues (Deeg & Verchere, 1997). The relevance of these observation to the present invention can be investigated using streptozotocin-induced diabetic rats. Streptozotcin results in destruction of the islets of Langerhans and thus these animals cannot produce

insulin and represent an animal model of type I diabetes mellitus (IDDM).

GPI-PLD levels can first be compared in the serum of streptozotocin-treated and control rats over time. If levels are consistently lower following streptozotocin treatment this would imply that the islets of Langerhans are indeed the major source of GPI-PLD found in serum. The ability of streptozotocin-treated and control rats to clear glucose following administration of either insulin alone or insulin combined with varying doses of GPI-PLD can also be examined. The GPI-PLD will be supplied bound to apolipoprotein A1 (apo A1) as it is in normal serum. If the addition of GPI-PLD increased the ability of insulin to reduce blood glucose, then this will have important implications in the treatment of diabetic patients.

Site of action

The function of the enzyme in releasing GPI-anchored proteins, and its postulated function as the generator of IPG second messengers require the enzyme to be active at the cell surface. It is known that GPI-anchored proteins accumulate in clusters in caveolae, an uncoated pit membrane specialisation, and so this is a good potential site for GPI-PLD activity. Analysis of the primary structure of the protein predicts a secondary structural arrangement of four amphipathic helices, thus suggesting that the protein can interact with lipids in membranes. Previous experiments have demonstrated significant amounts of the enzyme in the lyososmal fraction but not in the cytosol. The location of GPI-PLD will be examined by staining tissues with anti-GPI-PLD antibodies, followed by a gold particle-labelled second antibody. Tissue can then be prepared for transmission electron microscopy and the location of the GPI-PLD protein determined. Caveolae will also be produced according to

the protocol of Chang et al (1994), which involves three rounds of sucrose step gradient ultracentrifugation. Caveolae-enriched proteins will then be separated by SDS-PAGE and electrophorectically transferred to nitrocellulose membranes. We can then use the anti-GPI-PLD antibody to determine if GPI-PLD is present in these membrane specialisations.

Activation of GPI-PLD

If GPI-PLD is found to be phosphorylated by protein kinase C and/or protein kinase ck2 by MALDI-TOF spectrometry, the interaction of these proteins can be confirmed using immunoprecipitation since antibodies to GPI-PLD, protein kinase C and protein kinase ck2 have all be produced. The yeast two hybrid system can also be used to identify other proteins which interact with GPI-PLD in the cell. The yeast two hybrid systems (Chen et al, 1991) is based on the property of the yeast transcriptional activator Ga14, which is separable into DNA binding and transcriptional activating domains. GPI-PLD cDNAs can be cloned in frame into the DNA binding domain vector. This will be co-transfected into an appropriate yeast host strain along with a library of cDNAs cloned into the activation domain vector. Interaction of a protein with GPI-PLD will therefore result in localisation of the activation and DNA binding domains, and hence transcription of the galactosidase reporter gene. Clones containing interacting proteins are then identified by the colour reaction they produce. The advantage of this system is that the gene encoding the interacting protein is immediately available for sequence analysis and thus identification. The use of this system has enabled identification of many interacting proteins and the system available in kit form from Clontech. This also provides a method of screening for sustances which are capable of activating GPI-PLD, e.g. for further development as lead compounds.

Discussion

GPI-PLD is a metalloenzyme with 5 and 10 atoms per molecule of calcium and zinc, respectively. It circulates in a complex with apolipoprotein A1. GPI-PLD is produced in the pancreas by both α and β-cells in the islets of Langerhans. It is also produced by a mouse insulinoma cell line (TC3), with GPI-PLD and insulin generally colocalised intracellularly. The enzyme was shown to be secreted in response to insulin secretagogues. Both isoenzymes of GPI-PLD also seem to be present in liver; a major part of the activity could be washed away from the tissue by extraction with detergent-free buffer (thus, likely to be the plasma enzyme). There is some suggestions that the liver, as well as the pancreas, may contribute to the serum pool of GPI-PLD as patients with liver disease have lower levels of active enzyme, which is correlated with the reduced albumin levels.

It has been shown that streptozotcin-induced diabetes mellitus in the rat reduced the basal content of insulin-sensitive PIG in isolated hepatocytes by about 60%. The authors conclude that insulin resistance in these rats is related to the impairment of PIG metabolism. It has also been shown that the mRNA for a GPI-PLD-like gene was over expressed in genetically obese (ob/ob) mice in comparison to lean litter mates. In the context of the invention described herein, this latter finding suggests that GPI-PLD levels are responsive to the obese/diabetic genotype.

References:

The references mentioned herein are all incorporated by reference in their entirety.

Huang et al, US Patent No: 5,418,147.

Tsang et al, FASEB J. (supp), 6:A1922, 1992.

Scallon et al, Science, 252:446-448, 1991.

Hoener et al, Eur. J. Biochem., 206:747-757, 1992.

Li et al, J. Biol. Chem., 269:28963-28971, 1994.

Heller et al, Eur. J. Biochem., 224:823-833, 1994.

Jones et al, Biochem. Biophys. Res. Comm., 233:432-437, 1997.

Rademacher et al, Brazilian J. Med. Biol. Res., 27:327-341, 1994.

Lin et al, J. Cell Biol., 115:220a, 1991

Thompson et al, Nucleic Acid Research, 22:4673-4680, 1994, with algorithm from Higgins et al, CABIOS, 8(2):189-191, 1992.

Alemany et al, Nature, 330:77-79, 1987.

Caro et al, Biochem. Molec. Med., 61:214-228, 1997.

Deeg & Verchere, Endocrinology, 136:819-826, 1997.

Figure 1: Alignment of GPI-PLD deduced amino acid sequences

Top: protein produced from cDNA clone A1
Mid: protein produced from Roche patent bovine liver sequence
Bot: protein produced from Roche patent human liver sequence

```
MSAFRLWPGLLIMLG-SLCHRGSPCGLSTHIEIGHRALEFLQLHNGRVNYRELLLEHQDA
MSAFRFWSGLLMLLG-FLCPRSSPCGISTHIEIGHRALEFLHLQDGSINYKELLLRHQDA
MSAFRLWPGLLMIVMASLCHRGSSCGLSTHIEIGHRALEFLHLHNGHVNYKELLLEHQDA

YQAGIVFPDCFYPSICKGGKFHDVSESTHWTPFLNASVHYIRENYPLPWEKDTEKLVAFL
YQAGSVFPDSFYPSICERGQFHDVSESTHWTPFLNASVHYIRKNYPLPWDEDTEKLVAFL
YQAGTVFPDCFYPSLCKGGKFHDVSESTHWTPFLNASVHYIRENYPLPWEKDTEKLVAFL

FGITSHMAADVSWHSLGLEQGFLRTMGAIDFHGSYSEAHSAGDFGGDVLSQFEFNFNYLA
FGITSHMVADVNWHSLGIENGFLRTMAAIDFHNSYPEAHPAGDFGGDVLSQFEFKFNYLS
FGITSHMVADVSWHSLGIEQGFLRTMGAIDFHGSYSEAHSAGDFGGDVLSQFEFNFNYLA

RRWYVPVKDLLGIYEKLYGRKVITENVIVDCSHIQFLEMYGEMLAVSKLYPTYSTKSPFL
RHWYVPAEDLLGIYRELYGRIVITKKAIVDCSYLQFLEMYAEMLAISKLYPTYSVKSPFL
RRWYVPVKDLLGIYEKLYGREVITENVIVDCSHIQFLEMYGEMLAVSKLYPSYSTKSPFL

VEQFQEYFLGGLDDMAFWSTNIYHLTSFMLENGTSDCNLPENPENPLFIACGGQQNHTQG
VEQFQEYFLGGLEDMAFWSTNIYHLTSTMLKNGTSNCNLPENP---LFITCGGQQNNTHG
VEQFQEYFLGGLDDMAFWSTNIYHLTSFMLENGTSDCSLFENPENPLFIACGGQQNHTQG

SKMQKNDFHRNLTTSLTESVDRNINYTERGVFFSVNSWTPDSMSFIYKALERNIRTMFIG
SKVQKNGFHKNVTAALTKNIGKHINYTKRGVFFSVDSWTMDFLSFMYKSLERSIREMFIG
SKMQKNDFHRNLTSSLTENIDRNINYTERGVFFSVNSWTPDSMSFIYKALERNVRTMFIG

GSQLSQKHVSSPLASYFLSFPYARLGWAMTSADLNQDGHGDLVVGAPGYSRPGHIHIGRV
SSQP-LTHVSSPAASYYLSFPYTRLGWAMTSADLNQDGYGDLVVGAPGYSHPGRIHVGRV
GSQLSQKHISSPLASYFLSFPYARLGWAMTSADLNQDGYGDLVVGAPGYSRPGRIHIGRV

YLIYGNDLGLPPVDLDLDKEAHRILEGFQPSGRFGSALAVLDFNVDGVPDLAVGAPSVGS
YLIYGNDLG-PRIDLDLDKEAHGILEGFQPSGRFGSAVAVLDFNVDGVPDLAVGAPSVGS
YLIYGNELGLPPVDLDLDKEAHGILEGFQPSGRFGSALAMLDFNMDGVPDLAVGAPSVGS

EQLTYKGAVYVYFGSKQGGMSSSPNITISCQDIYCNLGWTLLAADVNGDSEPD-LVIGSP
EKLTYTGAVYVYFGSKQGQLSSSPNVTISCQDTYCNLGWTLLAADVDGDSEPDLFVIGSP
EQLTYKGAVYVYFGSKQGRMSSSPNITISCQDIYCNLGWTLLAADVNGDSEPD-LVIGSP

FAPGGGKQKGIVAAFYSGPSLSDKEKLNVEAANWTVRGEEDFSWFGYSLHGVTVDNRTLL
FAFGGGKQKGIVAAFYSGSSYSSREKLNVEAANWMVKGEEDFAWLGYSLHGVNVNNRTLL
FAPGGGKQKGIVAAFYSGPSLSNKEKLNVEAANWTVRGEEDFAWFGYSLHGVTVDNRTLL

LVGSPTWKNASRLGHLLHIRDEKKSLGRVYGYFPPNGQSWFTISGDKAMGKLGTSLSSGH
LAGSPTWKDTSSQGHLFRTRDEKQSPGRVYGYFPPICQSWFTISGDKAMGKLGTSLSSGH
LVGSPTWKNASRLGRLLHIRDEKKSLGRVYGYFPPNSQSWFTIVGDKAMGKLGTSLSSGH

VLMNGTLKQVLLVGAPTYDDVSKVAFLTVTLHQGGATRMYALISDAQPLLLSTFSGDRRF
VIVNGTRTQVLLVGAPTQDVVSKS-FLTMTLHQGGSTRMYELTPDSQPSLLSTFSGNRRF
VLMNGTLTQVLLVGAPTRDDVSKMAFLTMTLHQGGATRMYALTSDLQPPLLSTFSGDRRF

SRFGGVLHLSDLDDDGLDEIIMAAPLRIADVTSGLIGGEDGRVYVYNGKETTLGDMTGKC
SRFGGVLHLSDLDNDGLDEIIVAAPLRITDATAGLMGEEDGRVYVFNGKQITVGDVTGKC
SRFGGVLHLSDLDDDGVDEIIVAAPLRIADVTSGLIGGEDGRVYVYNGKETTLGDMTGKC

KSWITPCPEEKAQYVLISPEASSRFGSSLITVRSKAKNQVVIAAGRSSLGARLSGALHVY
KSWVTPCPEEKAQYVLISPEAGSRFGSSVITVRSKEKNQVIIAAGRSSLGARLSGVLHIY
KSWMTPCPEEKAQYVLISPEASSRFGSSLITVRSKAKNQVVIAAGRSSLGARLSGALHVY
```

SLGSD
RLGQD
SLGSD

Figure 2: Alignment of human GPI-PLD nucleic acid sequences

Top: pancreatic-form cDNA sequence from GenBank database
mid: our sequence cloned from human liver cDNA library
bot: Roche patent pancreatic-form partial cDNA sequence

```
      ------------------------------------------------------------
1     GTGACCTGCTTAGAGAGAAGCGGTGGGTCTGCACCTGGATTTTGGAGTCCCAGTGCTGCT   60
      ------------------------------------------------------------

1     ---------------------------------------------------ATGTCTGCT    9
61    GCAGCTCTGAGCATTCCCACGTCACCAGAGAAGCCGGTGGGCAATGAGAGCATGTCTGCT  120
      ------------------------------------------------------------

10    TTCAGGTTGTGGCCTGGCCTGCTGATCATGTTGGGTTCTCTCTGCCATAGAGGTTCACCG   69
121   TTCAGGTTGTGGCCTGGCCTGCTGATCATGTTGGGTTCTCTCTGCCATAGAGGTTCACCG  180
      ------------------------------------------------------------

70    TGTGGCCTTTCAACACACGTAGAAATAGGACACAGAGCTCTGGAGTTTCTTCAGCTTCAC  129
181   TGTGGCCTTTCAACACACATAGAAATAGGACACAGAGCTCTGGAGTTTCTTCAGCTTCAC  240
      ------------------------------------------------------------

130   AATGGGCGTGTTAACTACAGAGAGCTGTTACTAGAACACCAGGATGCGTATCAGGCTGGA  189
241   AATGGGCGTGTTAACTACAGAGAGCTGTTACTAGAACACCAGGATGCGTATCAGGCTGGA  300
      ------------------------------------------------------------

190   ATCGTGTTTCCTGATTGTTTTTACCCTAGCATCTGCAAAGGAGGAAAATTCCATGATGTG  249
301   ATCGTGTTTCCTGATTGTTTTTACCCTAGCATCTGCAAAGGAGGAAAATTCCATGATGTG  360
      ------------------------------------------------------------

250   TCTGAGAGCACTCACTGGACTCCGTTTCTTAATGCAAGCGTTCATTATATCCGAGAGAAC  309
361   TCTGAGAGCACTCACTGGACTCCGTTTCTTAATGCAAGCGTTCATTATATCCGAGAGAAC  420
      ------------------------------------------------------------

310   TATCCCCTTCCCTGGGAGAAGGACACAGAGAAACTGGTAGCTTTCTTGTTTGGAATTACT  369
421   TATCCCCTTCCCTGGGAGAAGGACACAGAGAAACTGGTAGCTTTCTTGTTTGGAATTACT  480
      ------------------------------------------------------------

370   TCTCACATGGCGGCAGATGTCAGCTGGCATAGTCTGGGCCTTGAACAAGGATTCCTTAGG  429
481   TCTCACATGGCGGCAGATGTCAGCTGGCATAGTCTGGGCCTTGAACAAGGATTCCTTAGG  540
      ------------------------------------------------------------

430   ACCATGGGAGCTATTGATTTTCACGGCTCCTATTCAGAGGCTCATTCGGCTGGTGATTTT  489
541   ACCATGGGAGCTATTGATTTTCACGGCTCCTATTCAGAGGCTCATTCGGCTGGTGATTTT  600
      ------------------------------------------------------------

490   GGAGGAGATGTGTTGAGCCAGTTTGAATTTAATTTTAATTACCTTGCACGACGCTGGTAT  549
601   GGAGGAGATGTGTTGAGCCAGTTTGAATTTAATTTTAATTACCTTGCACGACGCTGGTAT  660
      ------------------------------------------------------------

550   GTGCCAGTCAAAGATCTACTGGGAATTTATGAGAAACTGTATGGTCGAAAAGTCATCACC  609
661   GTGCCAGTCAAAGATCTACTGGGAATTTATGAGAAACTGTATGGTCGAAAAGTCATCACC  720
      ------------------------------------------------------------

610   GAAAATGTAATCGTTGATTGTTCACATATCCAGTTCTTAGAAATGTATGGTGAGATGCTA  669
721   GAAAATGTAATCGTTGATTGTTCACATATCCAGTTCTTAGAAATGTATGGTGAGATGCTA  780
      ------------------------------------------------------------

670   GCTGTTTCCAAGTTATATCCCACTTACTCTACAAAGTCCCCGTTTTTGGTGGAACAATTC  729
781   GCTGTTTCCAAGTTATATCCCACTTACTCTACAAAGTCCCCGTTTTTGGTGGAACAATTC  840
      ------------------------------------------------------------
```

```
730   CAAGAGTATTTTCTTGGAGGACTGGATGATATGGCATTTTGGTCCACTAATATTTACCAT   789
841   CAAGAGTATTTTCTTGGAGGACTGGATGATATGGCATTTTGGTCCACTAATATTTACCAT   900
      ------------------------------------------------------------

790   CTAACAATCTTCATGTTGGAGAATGGGACCAGTGACTGCAACCTGCCTGAGAACCCTCTG   849
901   CTAACAAGCTTCATGTTGGAGAATGGGACCAGTGACTGCAACCTGCCTGAGAACCCTCTG   960
      ------------------------------------------------------------

850   TTCATTGCATGTGGCGGCCAGCAAAACCACACCCAGGGCTCAAAAATGCAGAAAAATGAT   909
961   TTCATTGCATGTGGCGGCCAGCAAAACCACACCCAGGGCTCAAAAATGCAGAAAAATGAT  1020
      ------------------------------------------------------------

910   TTTCACAGAAATTTGACTACATCCCTAACTGAAAGTGTTGACAGGAATATAAACTATACT   969
1021  TTTCACAGAAATTTGACTACATCCCTAACTGAAAGTGTTGACAGGAATATAAACTATACT  1080
      ------------------------------------------------------------

970   GAAAGAGGAGTGTTCTTTAGTGTAAATTCCTGGACCCCGGATTCCATGTCCTTTATCTAC  1029
1081  GAAAGAGGAGTGTTCTTTAGTGTAAATTCCTGGACCCCGGATTCCATGTCCTTTATCTAC  1140
      ------------------------------------------------------------

1030  AAGGCTTTGGAAAGGAACATAAGGACAATGTTCATAGGTGGCTCTCAGTTGTCACAAAAG  1089
1141  AAGGCTTTGGAAAGGAACATAAGGACAATGTTCATAGGTGGCTCTCAGTTGTCACAAAAG  1200
      ------------------------------------------------------------

1090  CACGTCTCCAGCCCCTTAGCATCTTACTTCTTGTCATTTCCTTATGCGAGGCTTGGCTGG  1149
1201  CACGTCTCCAGCCCCTTAGCATCTTACTTCTTGTCATTTCCTTATGCGAGGCTTGGCTGG  1260
      ------------------------------------------------------------

1150  GCAATGACCTCAGCTGACCTCAACCAGGATGGGCACGGTGACCTCGTGGTGGGCGCACCA  1209
1261  GCAATGACCTCAGCTGACCTCAACCAGGATGGGCACGGTGACCTCGTGGTGGGCGCACCA  1320
      ------------------------------------------------------------

1210  GGCTACAGCCGCCCCGGCCACATCCACATCGGGCGCGTGTACCTCATCTACGGCAATGAC  1269
1321  GGCTACAGCCGCCCCGGCCACATCCACATCGGGCGCGTGTACCTCATCTACGGCAATGAC  1380
      ------------------------------------------------------------

1270  CTGGGCCTGCCACCTGTTGACCTGGACCTGGACAAGGAGGCCCACAGGATCCTTGAAGGC  1329
1381  CTGGGCCTGCCACCTGTTGACCTGGACCTGGACAAGGAGGCCCACAGGATCCTTGAAGGC  1440
      ------------------------------------------------------------

1330  TTCCAGCCCTCAGGTCGGTTTGGCTCGGCCTTGGCTGTGTTGGACTTTAACGTGGACGGC  1389
1441  TTCCAGCCCTCAGGTCGGTTTGGCTCGGCCTTGGCTGTGTTGGACTTTAACGTGGACGGC  1500
      ------------------------------------------------------------

1390  GTGCCTGACCTGGCCGTGGGAGCTCCCTCGGTGGGCTCCGAGCAGCTCACCTACAAAGGT  1449
1501  GTGCCTGACCTGGCCGTGGGAGCTCCCTCGGTGGGCTCCGAGCAGCTCACCTACAAAGGT  1560
      ------------------------------------------------------------

1450  GCCGTGTATGTCTACTTTGGTTCCAAACAAGGAGGAATGTCTTCTTCCCCTAACATCACC  1509
1561  GCCGTGTATGTCTACTTTGGTTCCAAACAAGGAGGAATGTCTTCTTCCCCTAACATCACC  1620
      ------------------------------------------------------------

1510  ATTTCTTGCCAGGACATCTACTGTAACTTGGGCTGGACTCTCTTGGCTGCAGATGTGAAT  1569
1621  ATTTCTTGCCAGGACATCTACTGTAACTTGGGCTGGACTCTCTTGGCTGCAGATGTGAAT  1680
      ------------------------------------------------------------

1570  GGAGACAGTGAACCCGATCTGGTCATCGGCTCCCCTTTTGCACCAGGTGGAGGGAAGCAG  1629
1681  GGAGACAGTGAACCCGATCTGGTCATCGGCTCCCCTTTTGCACCAGGTGGAGGGAAGCAG  1740
      ------------------------------------------------------------

1630  AAGGGAATTGTGGCTGCGTTTATTCTGGCCCCAGCCTGAGCGACAAAGAAAAACTGAAC  1689
1741  AAGGGAATTGTGGCTGCGTTTATTCTGGCCCCAGCCTGAGCGACAAAGAAAAACTGAAC  1800
1     -------------------------CTGGCCCCAGCCTGAGCGACAAAGAAAAACTGAAC    35
```

```
1690 GTGGAGGCAGCCAACTGGACGGTGAGAGGCGAGGAAGACTTCTCCTGGTTTGGATATTCC 1749
1801 GTGGAGGCAGCCAACTGGACGGTGAGAGGCGAGGAAGACTTCTCCTGGTTTGGATATTCC 1860
36   GTGGAGGCAGCCAACTGGACGGTGAGAGGCGAGGAAGACTTCTCCTGGTTTGGATATTCC 95

1750 CTTCACGGTGTCACTGTGGACAACAGAACCTTGCTGTTGGTTGGGAGCCCGACCTGGAAG 1809
1861 CTTCACGGTGTCACTGTGGACAACAGAACCTTGCTGTTGGTTGGGAGCCCGACCTGGAAG 1920
96   CTTCACGGTGTCACTGTGGACAACAGAACCTTGCTGTTGGTTGGGAGCCCGACCTGGAAG 155

1810 AATGCCAGCAGGCTGGGCCATTTGTTACACATCCGAGATGAGAAAAAGAGCCTTGGGAGG 1869
1921 AATGCCAGCAGGCTGGGCCATTTGTTACACATCCGAGATGAGAAAAAGAGCCTTGGGAGG 1980
156  AATGCCAGCAGGCTGGGCCATTTGTTACACATCCGAGATGAGAAAAAGAGCCTTGGGAGG 215

1870 GTGTATGGCTACTTCCCACCAAACGGCCAAAGCTGGTTTACCATTTCTGGAGACAAGGCA 1929
1981 GTGTATGGCTACTTCCCACCAAACGGCCAAAGCTGGTTTACCATTTCTGGAGACAAGGCA 2040
216  GTGTATGGCTACTTCC-ACCAAACGGCCAAAGCTGGTTTACCATTTCTGGAGACAAGGCA 275

1930 ATGGGGAAACTGGGTACTTCCCTTTCCAGTGGCCACGTACTGATGAATGGGACTCTGAAA 1989
2041 ATGGGGAAACTGGGTACTTCCCTTTCCAGTGGCCACGTACTGATGAATGGGACTCTGAAA 2100
276  ATGGGGAAACTGGGTACTTCCCTTTCCAGTGGTCACGTACTGATGAATGGGACTCTGAAA 335

1990 CAAGTGCTGCTGGTTGGAGCCCCTACGTACGATGACGTGTCTAAGGTGGCATTCCTGACC 2049
2101 CAAGTGCTGCTGGTTGGAGCCCCTACGTACGATGACGTGTCTAAGGTGGCATTCCTGACC 2160
336  CAAGTGCTGCTGGTTGGAGCCCCTACGTACGATGACGTGTCTAAGGTGGCATTCCTGACC 395

2050 GTGACCCTACACCAAGGCGGAGCCACTCGCATGTACGCACTCATATCTGACGCGCAGCCT 2109
2161 GTGACCCTACACCAAGGCGGAGCCACTCGCATGTACGCACTCATATCTGACGCGCAGCCT 2220
396  GTGACCCTACACCAAGGCGGAGCCACTCGCGTGTACGCACTCATATCTGACGCGCAGCCT 455

2110 CTGCTGCTCAGCACCTTCAGCGGAGACCGCCGCTTCTCCCGATTTGGTGGCGTTCTGCAC 2169
2221 CTGCTGCTCAGCACCTTCAGCGGAGACCGCCGCTTCTCCCGATTTGGTGGCGTTCTGCAC 2280
456  CTGCTGCTCAGCACCTTCAGCGGAGACCGCCGCTTCTCCCGATTTGGTGGCGTTCTGCAC 515

2170 TTGAGTGACCTGGATGATGATGGCTTAGATGAAATCATCATGGCAGCCCCCCTGAGGATA 2229
2281 TTGAGTGACCTGGATGATGATGGCTTAGATGAAATCATCATGGCAGCCCCCCTGAGGATA 2340
516  TTGAGTGACCTGGATGATGATGGCTTAGATGAAATCATCATGGCAGCCCCCCTGAGGATA 575

2230 GCAGATGTAACCTCTGGACTGATTGGGGGAGAAGACGGCCGAGTATATGTATATAATGGC 2289
2341 GCAGATGTAACCTCTGGACTGATTGGGGGAGAAGACGGCCGAGTATATGTATATAATGGC 2400
576  GCAGATGTAACCTCTGGACTGATTGGGGGAGAAGACGGCCGAGTATATGTATATAATGGC 635

2290 AAAGAGACCACCCTTGGTGACATGACTGGCAAATGCAAATCATGGATAACTCCATGTCCA 2349
2401 AAAGAGACCACCCTTGGTGACATGACTGGCAAATGCAAATCATGGATAACTCCATGTCCA 2460
636  AAAGAGACCACCCTTGGTGACATGACTGGCAAATGCAAATCATGGATAACTCCATGTCCA 695

2350 GAAGAAAAGGCCCAATATGTATTGATTTCTCCTGAAGCCAGCTCAAGGTTTGGGAGCTCC 2409
2461 GAAGAAAAGGCCCAATATGTATTGATTTCTCCTGAAGCCAGCTCAAGGTTTGGGAGCTCC 2520
696  GAAGAAAAGGCGCAATATGTATTGATTTCTCCTGAAGCCAGCTCAAGGTTTGGGAGCTCC 755

2410 CTCATCACCGTGAGGTCCAAGGCAAAGAACCAAGTCGTCATTGCTGCTGGAAGGAGTTCT 2469
2521 CTCATCACCGTGAGGTCCAAGGCAAAGAACCAAGTCGTCATTGCTGCTGGAAGGAGTTCT 2580
756  CTCATCACCGTGAGGTCCAAGGCAAAGAACCAAGTCGTCATTGCTGCTGGAAGGAGTTCT 815

2470 TTGGGAGCCCGACTCTCCGGGGCACTTCACGTCTATAGCCTTGGCTCAGATTGAAGATTT 2529
2581 TTGGGAGCCCGACTCTCCGGGGCACTTCACGTCTATAGCCTTGGCTCAGATTGAAGATTT 2640
816  TTGGGAGCCCGACTCTCCGGGGCACTTCACGTCTATAGCCTTGGCTCAGATTGAAGATTT 875

2530 CACTGCATTTCCCCACTCTGCCCACCTCTCTCATGCTGAATCACATCCATGGTGAGCATT 2589
2641 CACTGCATTTCCCCACTCTGCCCACCTCTCTCATGCTGAATCACATCCATGGTGAGCATT 2700
876  CACTGCATTTCCCCACTCTGCCCACCTCTCTCATGCTGAATCACATCCATGGTGAGCATT 935

2590 TTGATGGACAAAGTGGCACATCCAGTGGAGCGGTGGTAGATCCTGATAGACATGGGGCTC 2649
2701 TTGATGGACAAAGTGGCACATCCAGTGGAGCGGTGGTAGATCCTGATAGACATGGGGCTC 2760
936  TTGATGGACAAAGTGGCACATCCAGTGGAGCGGTGGTAGATCCTGATAGACATGGGGCTC 995
```

2.4

```
2650 CTGGGA------------------------------------------------------ 2655
2761 CTGGGA------------------------------------------------------ 2766
996  CTGGGACAGTGAACCCGATCTGGTCATCGGCTCCCCTTTTGCACCAGGTGGAGGGAAGCA 1055

2656 ----------------------------GTAGAGAGACACACTAACAGCCACACCCTCTG 2687
2767 ----------------------------GTAGAGAGACACACTAACAGCCACACCCTCTG 2798
1056 GAAGGGAATTGTGGCTGCGTTTTATTGAGTAGAGAGACACACTAACAGCCACACCCTCTG 1115

2688 GAAATCTGATACAGTAAATATATGACTGCACCAGAAATATGTGAAATAGCAGACATTCTG 2747
2799 GAAATCTGATACAGTAAATATATGACTGCACCAG-------------------------- 2833
1116 GAAATCTGATACAGTAAATATATGACTACACCAGAAATATGTGAAATAGCAGACATTCTG 1175

2748 CTTACTCATGTCTCCTTCCACAGTTTACTTCCTCGCTCCCTTTGCATCTAAACCTTTCTT 2807
     ------------------------------------------------------------
1176 CTTACTCATGTCTCCTTCCACAGTTTACTTCCTCGCTCCCTTTGCATCTAAACCTTTCTT 1235

2808 CTTTCCCAACTTATTGCCTGTAGTCAGACCTGCTGTACAACCTATTTCCTCTTCCTCTTG 2867
     ------------------------------------------------------------
1236 CTTTCCCAACTTATTGCCTGTAGTC----------------------------------- 1261

2868 AATGTCTTTCCAGTGGCTGGAAAGGTCCCTCTGTGGTTATCTGTTAGAACAGTCTCTGTA 2927
     ------------------------------------------------------------
     ------------------------------------------------------------

2928 CACAATTCCTCCTAAAAACATCCTTTTTTAAAAAAAGAATTGTTCAGCCATAAAGAAAGA 2987
     ------------------------------------------------------------
     ------------------------------------------------------------

2988 ACAAGATCATGCCCTTTGCAGGGACATGGATGGAGCTGGAGGCCATTATCCTTCATAAAC 3047
     ------------------------------------------------------------
     ------------------------------------------------------------

3048 TATTGCAGGAACAGAAAACCAAACACTCCATATTCTCACTTGTAAGTGGGAGCTAAGTGA 3107
     ------------------------------------------------------------
     ------------------------------------------------------------

3108 GAACACGTGGACACATAGAGGGAAACAACACACACTGGGGCCTATGAGAGGGCGGAAGGT 3167
     ------------------------------------------------------------
     ------------------------------------------------------------

3168 GGGAGGAGGGAGAGATCAGGAAAAATAACTAATGGATACTTAGGGTGATGAAATAATCTG 3227
     ------------------------------------------------------------
     ------------------------------------------------------------

3228 TGTAACAAACCCCCATGACACACCTTTATGTATGTAACAAACCAGCACTTCCTGCGCATG 3287
     ------------------------------------------------------------
     ------------------------------------------------------------

3288 TACCCCTGAACTTAAAAGTTAAAAAAAAAGTTGAACTTAAAAATAACAGATTGGCCCATGC 3347
     ------------------------------------------------------------
     ------------------------------------------------------------

3348 CAATCAAAGTATAATAGAAAGCATAGTATAC 3378
     -------------------------------
     -------------------------------
```

Figure 3: 'Amino acid sequences of GPI-PLD a1, b2 and d3.

cDNA clone d3

MILLFQDSMSFIYKALERNIRTMFIGGSQLSQKHVSSPLASYFLSFPYARLGWAMTSADL
NQDGHGDLVVGAPGYSRPGHIHIGRVYLIYGNDLGLPPVDLDLDKEAHRILEGFQPSGRF
GSALAVLDFNVDGVPDLAVGAPSVGSEQLTKGAVYVYFGSKQGGMSSSPNITISCQDIYC
NLGWTLLAADVNGDSEPDLVIGSPFAPGGGKQKGIVAAFYSGPSLSDKEKLNVEAANWTV
RGEEDFSWFGYSLHGVTVDNRTLLLVGSPTWKNASRLGHLLHIRDEKKSLGRVYGYFPPN
GQSWFTISGDKAMGKLGTSLSSGHVLMNGTLKQVLLVGAPTYDDVSKVAFLTVTLHQGGA
TRMYALISDAQPLLLSTFSGDRRFSRFGGVLHLSDLDDDGLDEIIMAAPLRIADVTSGLI
GGEDGRVYVYNGKETTLGDMTGKCKSWITPCPEEKAQYVLISPEASSRFGSSLITVRSKA
KNQVVIAAGRSSLGARLSGALHVYSLGSD

cDNA clone b2

MSAFRLWPGLLIMLGSLCHRGSPCGLSTHIEIGHRALEFLQLHNGRVNYRELLLEHQDAY
QAGIVFPDCFYPSICKGGKFHDVSESTHWTPFLNASVHYIRENYPLPWEKDTEKLVAFLF
GITSHMAADVSWHSLGLEQGFLRTMGAIDFHGSYSEAHSAGDFGGDVLSQFEFNFNYLAR
RWYVPVKDLLGIYEKLYGRKVITENVIVDCSHIQFLEMYGEMLAVSKLYPTYSTKSPFLV
EQFQEYFLGGLDDMAFWSTNIYHLTSFMLENGTSDCNLPENPLFIACGGQQNHTQGSKMQ
KNDFHRNLTTSLTESVDRNINYTERGVFFSVNSWTPDSMSFIYKALERNIRTMFIGGSQL
SQKHVSSPLASYFLSFPYARLGWAMTSADLNQDGHGDLVVGAPGYSRPGHIHIGRVYLIY
GNDLGLPPVDLDLDKEAHRILEGFQPSGRFGSALAVLDFNVDGVPDLAVGAPSVGSEQLT
YKGAVYVYFGSKQGGMSSSPNITISCQDIYCNLGWTLLAADVNGDSEPDLVIGSPFAPGG
GKQKGIVAAFYSGPSLSDKEKLNVEAANWTVRGEEDFSWFGYSLHGVTVDNRTLLLVGSP
TWKNASRLGHLLHIRDEKKSLGRVYGYFPPNGQSWFTISGDKAMGKLGTSLSSGHVLMNG
TLKQVLLVGAPTYDDVSKVAFLTVTLHQGGATRMYALISDAQPLLLSTFSGDRRFSRFGG
VLHLSDLDDDGLDEIIMAAPLRIADVTSGLIGGEDGRVYVYNGKETTLGDMTGKCKSWIT
PCPEEKVSEKKKKKK

cDNA clone a1

MSAFRLWPGLLIMLGSLCHRGSPCGLSTHIEIGHRALEFLQLHNGRVNYRELLLEHQDAY
QAGIVFPDCFYPSICKGGKFHDVSESTHWTPFLNASVHYIRENYPLPWEKDTEKLVAFLF
GITSHMAADVSWHSLGLEQGFLRTMGAIDFHGSYSEAHSAGDFGGDVLSQFEFNFNYLAR
RWYVPVKDLLGIYEKLYGRKVITENVIVDCSHIQFLEMYGEMLAVSKLYPTYSTKSPFLV
EQFQEYFLGGLDDMAFWSTNIYHLTSFMLENGTSDCNLPENPLFIACGGQQNHTQGSKMQ
KNDFHRNLTTSLTESVDRNINYTERGVFFSVNSWTPDSMSFIYKALERNIRTMFIGGSQL
SQKHVSSPLASYFLSFPYARLGWAMTSADLNQDGHGDLVVGAPGYSRPGHIHIGRVYLIY
GNDLGLPPVDLDLDKEAHRILEGFQPSGRFGSALAVLDFNVDGVPDLAVGAPSVGSEQLT
YKGAVYVYFGSKQGGMSSSPNITISCQDIYCNLGWTLLAADVNGDSEPDLVIGSPFAPGG
GKQKGIVAAFYSGPSLSDKEKLNVEAANWTVRGEEDFSWFGYSLHGVTVDNRTLLLVGSP
TWKNASRLGHLLHIRDEKKSLGRVYGYFPPNGQSWFTISGDKAMGKLGTSLSSGHVLMNG
TLKQVLLVGAPTYDDVSKVAFLTVTLHQGGATRMYALISDAQPLLLSTFSGDRRFSRFGG
VLHLSDLDDDGLDEIIMAAPLRIADVTSGLIGGEDGRVYVYNGKETTLGDMTGKCKSWIT
PCPEEKAQYVLISPEASSRFGSSLITVRSKAKNQVVIAAGRSSLGARLSGALHVYSLGSD

Figure 4: Human GPI-PLD cDNA clone a1

2832 bp: 690 a 688 c 735 g 719 t

```
   1 gtgacctgct tagagagaag cggtgggtct gcacctggat tttggagtcc cagtgctgct
  61 gcagctctga gcattcccac gtcaccagag aagccggtgg gcaatgagag catgtctgct
 121 ttcaggttgt ggcctggcct gctgatcatg ttgggttctc tctgccatag aggttcaccg
 181 tgtggccttt caacacacat agaaatagga cacagagctc tggagtttct tcagcttcac
 241 aatgggcgtg ttaactacag agagctgtta ctagaacacc aggatgcgta tcaggctgga
 301 atcgtgtttc ctgattgttt ttaccctagc atctgcaaag gaggaaaatt ccatgatgtg
 361 tctgagagca ctcactggac tccgtttctt aatgcaagcg ttcattatat ccgagagaac
 421 tatccccttc cctgggagaa ggacacagag aaactggtag ctttcttgtt tggaattact
 481 tctcacatgg cggcagatgt cagctggcat agtctgggcc ttgaacaagg attccttagg
 541 accatgggag ctattgattt tcacggctcc tattcagagg ctcattcggc tggtgatttt
 601 ggaggagatg tgttgagcca gtttgaattt aattttaatt accttgcacg acgctggtat
 661 gtgccagtca aagatctact gggaatttat gagaaactgt atggtcgaaa agtcatcacc
 721 gaaaatgtaa tcgttgattg ttcacatatc cagttcttag aaatgtatgg tgagatgcta
 781 gctgtttcca agttatatcc cacttactct acaaagtccc cgtttttggt ggaacaattc
 841 caagagtatt ttcttggagg actggatgat atggcatttt ggtccactaa tatttaccat
 901 ctaacaagct tcatgttgga gaatgggacc agtgactgca acctgcctga gaaccctctg
 961 ttcattgcat gtggcggcca gcaaaaccac acccagggct caaaaatgca gaaaaatgat
1021 tttcacagaa atttgactac atccctaact gaaagtgttg acaggaatat aaactatact
1081 gaaagaggag tgttctttag tgtaaattcc tggaccccgg attccatgtc ctttatctac
1141 aaggctttgg aaaggaacat aaggacaatg ttcataggtg gctctcagtt gtcacaaaag
1201 cacgtctcca gcccccttagc atcttacttc ttgtcatttc cttatgcgag gcttggctgg
1261 gcaatgacct cagctgacct caaccaggat gggcacggtg acctcgtggt gggcgcacca
1321 ggctacagcc gccccggcca catccacatc gggcgcgtgt acctcatcta cggcaatgac
1381 ctgggcctgc cacctgttga cctggacctg gacaaggagg cccacaggat ccttgaaggc
1441 ttccagccct caggtcggtt tggctcggcc ttggctgtgt tggactttaa cgtggacggc
1501 gtgcctgacc tggccgtggg agctccctcg gtgggctccg agcagctcac ctacaaaggt
1561 gccgtgtatg tctactttgg ttccaaacaa ggaggaatgt cttcttcccc taacatcacc
1621 atttcttgcc aggacatcta ctgtaacttg ggctggactc tcttggctgc agatgtgaat
1681 ggagacagtg aacccgatct ggtcatcggc tccccttttg caccaggtgg agggaagcag
1741 aagggaattg tggctgcgtt ttattctggc cccagcctga gcgacaaaga aaaactgaac
1801 gtggaggcag ccaactggac ggtgagaggc gaggaagact tctcctggtt tggatattcc
1861 cttcacggtg tcactgtgga caacagaacc ttgctgttgg ttgggagccc gacctggaag
1921 aatgccagca ggctgggcca tttgttacac atccgagatg agaaaaagag ccttgggagg
1981 gtgtatggct acttcccacc aaacggccaa agctggttta ccatttctgg agacaaggca
2041 atggggaaac tgggtacttc cctttccagt ggccacgtac tgatgaatgg gactctgaaa
2101 caagtgctgc tggttggagc ccctacgtac gatgacgtgt ctaaggtggc attcctgacc
2161 gtgaccctac accaaggcgg agccactcgc atgtacgcac tcatatctga cgcgcagcct
2221 ctgctgctca gcaccttcag cggagaccgc cgcttctccc gatttggtgg cgttctgcac
2281 ttgagtgacc tggatgatga tggcttagat gaaatcatca tggcagcccc cctgaggata
2341 gcagatgtaa cctctggact gattggggga gaagacggcc gagtatatgt atataatggc
2401 aaagagacca ccccttggtga catgactggc aaatgcaaat catggataac tccatgtcca
2461 gaagaaaagg cccaatatgt attgatttct cctgaagcca gctcaaggtt tgggagctcc
2521 ctcatcaccg tgaggtccaa ggcaaagaac caagtcgtca ttgctgctgg aaggagttct
2581 ttgggagccc gactctccgg ggcacttcac gtctatagcc ttggctcaga ttgaagattt
2641 cactgcattt ccccactctg cccacctctc tcatgctgaa tcacatccat ggtgagcatt
2701 ttgatggaca aagtggcaca tccagtggag cggtggtaga tcctgataga catggggctc
2761 ctgggagtag agagacacac taacagccac accctctgga aatctgatac agtaaatata
2821 tgactgcacc ag
```

Figure 5: Human GPI-PLD cDNA clone b2

2472 bp: 617 a 588 c 639 g 628 t

```
   1 gtctgcacct ggattttgga gtcccagtgc tgctgcagct ctgagcattc ccacgtcacc
  61 agagaagccg gtgggcaatg agagcatgtc tgctttcagg ttgtggcctg gcctgctgat
 121 catgttgggt tctctctgcc atagaggttc accgtgtggc ctttcaacac acatagaaat
 181 aggacacaga gctctggagt ttcttcagct tcacaatggg cgtgttaact acagagagct
 241 gttactagaa caccaggatg cgtatcaggc tggaatcgtg tttcctgatt gtttttaccc
 301 tagcatctgc aaaggaggaa aattccatga tgtgtctgag agcactcact ggactccgtt
 361 tcttaatgca agcgttcatt atatccgaga gaactatccc cttccctggg agaaggacac
 421 agagaaactg gtagctttct tgtttggaat tacttctcac atggcggcag atgtcagctg
 481 gcatagtctg ggccttgaac aaggattcct taggaccatg ggagctattg attttcacgg
 541 ctcctattca gaggctcatt cggctggtga ttttggagga gatgtgttga gccagtttga
 601 atttaatttt aattaccttg cacgacgctg gtatgtgcca gtcaaagatc tactgggaat
 661 ttatgagaaa ctgtatggtc gaaaagtcat caccgaaaat gtaatcgttg attgttcaca
 721 tatccagttc ttagaaatgt atggtgagat gctagctgtt tccaagttat atcccactta
 781 ctctacaaag tccccgtttt tggtggaaca attccaagag tattttcttg gaggactgga
 841 tgatatggca ttttggtcca ctaatattta ccatctaaca agcttcatgt tggagaatgg
 901 gaccagtgac tgcaacctgc ctgagaaccc tctgttcatt gcatgtggcg gccagcaaaa
 961 ccacacccag ggctcaaaaa tgcagaaaaa tgattttcac agaaatttga ctacatccct
1021 aactgaaagt gttgacagga atataaacta tactgaaaga ggagtgttct ttagtgtaaa
1081 ttcctggacc ccggattcca tgtcctttat ctacaaggct ttggaaagga acataaggac
1141 aatgttcata ggtggctctc agttgtcaca aaagcacgtc tccagcccct tagcatctta
1201 cttcttgtca tttccttatg cgaggcttgg ctgggcaatg acctcagctg acctcaacca
1261 ggatgggcac ggtgacctcg tggtgggcgc accaggctac agccgccccg gccacatcca
1321 catcgggcgc gtgtacctca tctacggcaa tgacctgggc ctgccacctg ttgacctgga
1381 cctggacaag gaggcccaca ggatccttga aggcttccag ccctcaggtc ggtttggctc
1441 ggccttggct gtgttggact ttaacgtgga cggcgtgcct gacctggccg tgggagctcc
1501 ctcggtgggc tccgagcagc tcacctacaa aggtgccgtg tatgtctact ttggttccaa
1561 acaaggagga atgtcttctt cccctaacat caccatttct tgccaggaca tctactgtaa
1621 cttgggctgg actctcttgg ctgcagatgt gaatggagac agtgaacccg atctggtcat
1681 cggctcccct tttgcaccag gtggagggaa gcagaaggga attgtggctg cgttttattc
1741 tggccccagc ctgagcgaca aagaaaaact gaacgtggag gcagccaact ggacggtgag
1801 aggcgaggaa gacttctcct ggtttggata ttcccttcac ggtgtcactg tggacaacag
1861 aaccttgctg ttggttggga gcccgacctg gaagaatgcc agcaggctgg gccatttgtt
1921 acacatccga gatgagaaaa agagccttgg gagggtgtat ggctacttcc caccaaacgg
1981 ccaaagctgg tttaccattt ctggagacaa ggcaatgggg aaactgggta cttccctttc
2041 cagtggccac gtactgatga atgggactct gaaacaagtg ctgctggttg gagcccctac
2101 gtacgatgac gtgtctaagg tggcattcct gaccgtgacc ctacaccaag gcggagccac
2161 tcgcatgtac gcactcatat ctgacgcgca gcctctgctg ctcagcacct tcagcggaga
2221 ccgccgcttc tcccgatttg gtggcgttct gcacttgagt gacctggatg atgatggctt
2281 agatgaaatc atcatggcag ccccccctgag gatagcagat gtaacctctg gactgattgg
2341 gggagaagac ggccgagtat atgtatataa tggcaaagag accacccttg gtgacatgac
2401 tggcaaatgc aaatcatgga taactccatg tccagaagaa aaggtaagtg aaaaaaaaaa
2461 aaaaaaaaaa aa
```

Figure 6: Human GPI-PLD cDNA clone d3

1942 bp: 455 a 496 c 502 g 489 t

```
   1 gggctgtaac tctgccatcc ctcagcataa tttgggggta tgatttcact atcctaattg
  61 cctgtcctaa gtgatcttac ttgctgatag gacctaatgt tttattttat tgtttagcac
 121 ttctaaaaac tcatttcctt tacacaagtc caatactttg gacaggaaac agtagctttg
 181 ttgattatgc tacgtgtctt tactgtctat aatgattctt ttatttcagg attccatgtc
 241 ctttatctac aaggctttgg aaaggaacat aaggacaatg ttcataggtg gctctcagtt
 301 gtcacaaaag cacgtctcca gccccttagc atcttacttc ttgtcatttc cttatgcgag
 361 gcttggctgg gcaatgacct cagctgacct caaccaggat gggcacggtg acctcgtggt
 421 gggcgcacca ggctacagcc gccccggcca catccacatc gggcgcgtgt acctcatcta
 481 cggcaatgac ctgggcctgc cacctgttga cctggacctg gacaaggagg cccacaggat
 541 ccttgaaggc ttccagccct caggtcggtt tggctcggcc ttggctgtgt tggactttaa
 601 cgtggacggc gtgcctgacc tggccgtggg agctccctcg gtgggctccg agcagctcac
 661 ctacaaaggt gccgtgtatg tctactttgg ttccaaacaa ggaggaatgt cttcttcccc
 721 taacatcacc atttcttgcc aggacatcta ctgtaacttg ggctggactc tcttggctgc
 781 agatgtgaat ggagacagtg aacccgatct ggtcatcggc tccccttttg caccaggtgg
 841 agggaagcag aagggaattg tggctgcgtt ttattctggc cccagcctga gcgacaaaga
 901 aaaactgaac gtggaggcag ccaactggac ggtgagaggc gaggaagact tctcctggtt
 961 tggatattcc cttcacggtg tcactgtgga caacagaacc ttgctgttgg ttgggagccc
1021 gacctggaag aatgccagca ggctgggcca tttgttacac atccgagatg agaaaaagag
1081 ccttgggagg gtgtatggct acttcccacc aaacggccaa agctggttta ccatttctgg
1141 agacaaggca atggggaaac tgggtacttc cctttccagt ggccacgtac tgatgaatgg
1201 gactctgaaa caagtgctgc tggttggagc ccctacgtac gatgacgtgt ctaaggtggc
1261 attcctgacc gtgaccctac accaaggcgg agccactcgc atgtacgcac tcatatctga
1321 cgcgcagcct ctgctgctca gcaccttcag cggagaccgc cgcttctccc gatttggtgg
1381 cgttctgcac ttgagtgacc tggatgatga tggcttagat gaaatcatca tggcagcccc
1441 cctgaggata gcagatgtaa cctctggact gattggggga gaagacggcc gagtatatgt
1501 atataatggc aaagagacca cccttggtga catgactggc aaatgcaaat catggataac
1561 tccatgtcca gaagaaaagg cccaatatgt attgatttct cctgaagcca gctcaaggtt
1621 tgggagctcc ctcatcaccg tgaggtccaa ggcaaagaac caagtcgtca ttgctgctgg
1681 aaggagttct ttgggagccc gactctccgg ggcacttcac gtctatagcc ttggctcaga
1741 ttgaagattt cactgcattt ccccactctg cccacctctc tcatgctgaa tcacatccat
1801 ggtgagcatt ttgatggaca aagtggcaca tccagtggag cggtggtaga tcctgataga
1861 catggggctc ctgggagtag agagacacac taacagccac accctctgga aatctgatac
1921 agtaaatata tgactgcacc ag
```

Figure 7: Alignment of GPIPLD protein sequences

```
database    MSAFRLWPGLLIMLGSLCHRGSPCGLSTHVEIGHRALEFLQLHNGRVNYRELLLEHQDAY   60
d3          ------------------------------------------------------------
b2          MSAFRLWPGLLIMLGSLCHRGSPCGLSTHIEIGHRALEFLQLHNGRVNYRELLLEHQDAY   60
a1          MSAFRLWPGLLIMLGSLCHRGSPCGLSTHIEIGHRALEFLQLHNGRVNYRELLLEHQDAY   60

database    QAGIVFPDCFYPSICKGGKFHDVSESTHWTPFLNASVHYIRENYPLPWEKDTEKLVAFLF  120
d3          ------------------------------------------------------------
b2          QAGIVFPDCFYPSICKGGKFHDVSESTHWTPFLNASVHYIRENYPLPWEKDTEKLVAFLF  120
a1          QAGIVFPDCFYPSICKGGKFHDVSESTHWTPFLNASVHYIRENYPLPWEKDTEKLVAFLF  120

database    GITSHMAADVSWHSLGLEQGFLRTMGAIDFHGSYSEAHSAGDFGGDVLSQFEFNFNYLAR  180
d3          ------------------------------------------------------------
b2          GITSHMAADVSWHSLGLEQGFLRTMGAIDFHGSYSEAHSAGDFGGDVLSQFEFNFNYLAR  180
a1          GITSHMAADVSWHSLGLEQGFLRTMGAIDFHGSYSEAHSAGDFGGDVLSQFEFNFNYLAR  180

database    RWYVPVKDLLGIYEKLYGRKVITENVIVDCSHIQFLEMYGEMLAVSKLYPTYSTKSPFLV  240
d3          ------------------------------------------------------------
b2          RWYVPVKDLLGIYEKLYGRKVITENVIVDCSHIQFLEMYGEMLAVSKLYPTYSTKSPFLV  240
a1          RWYVPVKDLLGIYEKLYGRKVITENVIVDCSHIQFLEMYGEMLAVSKLYPTYSTKSPFLV  240

database    EQFQEYFLGGLDDMAFWSTNIYHLTIFMLENGTSDCNLPENPLFIACGGQQNHTQGSKMQ  300
d3          ------------------------------------------------------------
b2          EQFQEYFLGGLDDMAFWSTNIYHLTSFMLENGTSDCNLPENPLFIACGGQQNHTQGSKMQ  300
a1          EQFQEYFLGGLDDMAFWSTNIYHLTSFMLENGTSDCNLPENPLFIACGGQQNHTQGSKMQ  300

database    KNDFHRNLTTSLTESVDRNINYTERGVFFSVNSWTPDSMSFIYKALERNIRTMFIGGSQL  360
d3          ------------------------------MILLFQDSMSFIYKALERNIRTMFIGGSQL   30
b2          KNDFHRNLTTSLTESVDRNINYTERGVFFSVNSWTPDSMSFIYKALERNIRTMFIGGSQL  360
a1          KNDFHRNLTTSLTESVDRNINYTERGVFFSVNSWTPDSMSFIYKALERNIRTMFIGGSQL  360

database    SQKHVSSPLASYFLSFPYARLGWAMTSADLNQDGHGDLVVGAPGYSRPGHIHIGRVYLIY  420
d3          SQKHVSSPLASYFLSFPYARLGWAMTSADLNQDGHGDLVVGAPGYSRPGHIHIGRVYLIY   90
b2          SQKHVSSPLASYFLSFPYARLGWAMTSADLNQDGHGDLVVGAPGYSRPGHIHIGRVYLIY  420
a1          SQKHVSSPLASYFLSFPYARLGWAMTSADLNQDGHGDLVVGAPGYSRPGHIHIGRVYLIY  420

database    GNDLGLPPVDLDLDKEAHRILEGFQPSGRFGSALAVLDFNVDGVPDLAVGAPSVGSEQLT  480
d3          GNDLGLPPVDLDLDKEAHRILEGFQPSGRFGSALAVLDFNVDGVPDLAVGAPSVGSEQLT  150
b2          GNDLGLPPVDLDLDKEAHRILEGFQPSGRFGSALAVLDFNVDGVPDLAVGAPSVGSEQLT  480
a1          GNDLGLPPVDLDLDKEAHRILEGFQPSGRFGSALAVLDFNVDGVPDLAVGAPSVGSEQLT  480

database    YKGAVYVYFGSKQGGMSSSPNITISCQDIYCNLGWTLLAADVNGDSEPDLVIGSPFAPGG  540
d3          YKGAVYVYFGSKQGGMSSSPNITISCQDIYCNLGWTLLAADVNGDSEPDLVIGSPFAPGG  210
b2          YKGAVYVYFGSKQGGMSSSPNITISCQDIYCNLGWTLLAADVNGDSEPDLVIGSPFAPGG  540
a1          YKGAVYVYFGSKQGGMSSSPNITISCQDIYCNLGWTLLAADVNGDSEPDLVIGSPFAPGG  540

database    GKQKGIVAAFYSGPSLSDKEKLNVEAANWTVRGEEDFSWFGYSLHGVTVDNRTLLLVGSP  600
d3          GKQKGIVAAFYSGPSLSDKEKLNVEAANWTVRGEEDFSWFGYSLHGVTVDNRTLLLVGSP  270
b2          GKQKGIVAAFYSGPSLSDKEKLNVEAANWTVRGEEDFSWFGYSLHGVTVDNRTLLLVGSP  600
a1          GKQKGIVAAFYSGPSLSDKEKLNVEAANWTVRGEEDFSWFGYSLHGVTVDNRTLLLVGSP  600

database    TWKNASRLGHLLHIRDEKKSLGRVYGYFPPNGQSWFTISGDKAMGKLGTSLSSGHVLMNG  660
d3          TWKNASRLGHLLHIRDEKKSLGRVYGYFPPNGQSWFTISGDKAMGKLGTSLSSGHVLMNG  330
b2          TWKNASRLGHLLHIRDEKKSLGRVYGYFPPNGQSWFTISGDKAMGKLGTSLSSGHVLMNG  660
a1          TWKNASRLGHLLHIRDEKKSLGRVYGYFPPNGQSWFTISGDKAMGKLGTSLSSGHVLMNG  660

database    TLKQVLLVGAPTYDDVSKVAFLTVTLHQGGATRMYALISDAQPLLLSTFSGDRRFSRFGG  720
d3          TLKQVLLVGAPTYDDVSKVAFLTVTLHQGGATRMYALISDAQPLLLSTFSGDRRFSRFGG  390
b2          TLKQVLLVGAPTYDDVSKVAFLTVTLHQGGATRMYALISDAQPLLLSTFSGDRRFSRFGG  720
a1          TLKQVLLVGAPTYDDVSKVAFLTVTLHQGGATRMYALISDAQPLLLSTFSGDRRFSRFGG  720
```

```
database   VLHLSDLDDDGLDEIIMAAPLRIADVTSGLIGGEDGRVYVYNGKETTLGDMTGKCKSWIT  780
d3         VLHLSDLDDDGLDEIIMAAPLRIADVTSGLIGGEDGRVYVYNGKETTLGDMTGKCKSWIT  450
b2         VLHLSDLDDDGLDEIIMAAPLRIADVTSGLIGGEDGRVYVYNGKETTLGDMTGKCKSWIT  780
a1         VLHLSDLDDDGLDEIIMAAPLRIADVTSGLIGGEDGRVYVYNGKETTLGDMTGKCKSWIT  780

database   PCPEEKAQYVLISPEASSRFGSSLITVRSKAKNQVVIAAGRSSLGARLSGALHVYSLGSD  840
d3         PCPEEKAQYVLISPEASSRFGSSLITVRSKAKNQVVIAAGRSSLGARLSGALHVYSLGSD  510
b2         PCPEEKVSEKKKKKK---------------------------------------------  795
a1         PCPEEKAQYVLISPEASSRFGSSLITVRSKAKNQVVIAAGRSSLGARLSGALHVYSLGSD  840
```

Database	840 aa
d3	510 aa
b2	795 aa
a1	840 aa

Figure 8: Alignment of human GPI-PLD nucleic acid sequences

1: pancreatic-form: cDNA sequence from GenBank database (L11702)
2: cDNA clone A1
3: cDNA clone B2
4: cDNA clone D3

```
     ------------------------------------------------------------
1    GTGACCTGCTTAGAGAGAAGCGGTGGGTCTGCACCTGGATTTTGGAGTCCCAGTGCTGCT   60
1    --------------------------GTCTGCACCTGGATTTTGGAGTCCCAGTGCTGCT   34
     ------------------------------------------------------------

1    ---------------------------------------------------ATGTCTGCT    9
61   GCAGCTCTGAGCATTCCCACGTCACCAGAGAAGCCGGTGGGCAATGAGAGCATGTCTGCT  120
35   GCAGCTCTGAGCATTCCCACGTCACCAGAGAAGCCGGTGGGCAATGAGAGCATGTCTGCT   94
     ------------------------------------------------------------

10   TTCAGGTTGTGGCCTGGCCTGCTGATCATGTTGGGTTCTCTCTGCCATAGAGGTTCACCG   69
121  TTCAGGTTGTGGCCTGGCCTGCTGATCATGTTGGGTTCTCTCTGCCATAGAGGTTCACCG  180
95   TTCAGGTTGTGGCCTGGCCTGCTGATCATGTTGGGTTCTCTCTGCCATAGAGGTTCACCG  154
     ------------------------------------------------------------

70   TGTGGCCTTTCAACACACGTAGAAATAGGACACAGAGCTCTGGAGTTTCTTCAGCTTCAC  129
181  TGTGGCCTTTCAACACACATAGAAATAGGACACAGAGCTCTGGAGTTTCTTCAGCTTCAC  240
155  TGTGGCCTTTCAACACACATAGAAATAGGACACAGAGCTCTGGAGTTTCTTCAGCTTCAC  214
     ------------------------------------------------------------

130  AATGGGCGTGTTAACTACAGAGAGCTGTTACTAGAACACCAGGATGCGTATCAGGCTGGA  189
241  AATGGGCGTGTTAACTACAGAGAGCTGTTACTAGAACACCAGGATGCGTATCAGGCTGGA  300
215  AATGGGCGTGTTAACTACAGAGAGCTGTTACTAGAACACCAGGATGCGTATCAGGCTGGA  274
     ------------------------------------------------------------

190  ATCGTGTTTCCTGATTGTTTTTACCCTAGCATCTGCAAAGGAGGAAAATTCCATGATGTG  249
301  ATCGTGTTTCCTGATTGTTTTTACCCTAGCATCTGCAAAGGAGGAAAATTCCATGATGTG  360
275  ATCGTGTTTCCTGATTGTTTTTACCCTAGCATCTGCAAAGGAGGAAAATTCCATGATGTG  334
     ------------------------------------------------------------

250  TCTGAGAGCACTCACTGGACTCCGTTTCTTAATGCAAGCGTTCATTATATCCGAGAGAAC  309
361  TCTGAGAGCACTCACTGGACTCCGTTTCTTAATGCAAGCGTTCATTATATCCGAGAGAAC  420
335  TCTGAGAGCACTCACTGGACTCCGTTTCTTAATGCAAGCGTTCATTATATCCGAGAGAAC  394
     ------------------------------------------------------------

310  TATCCCCTTCCCTGGGAGAAGGACACAGAGAAACTGGTAGCTTTCTTGTTTGGAATTACT  369
421  TATCCCCTTCCCTGGGAGAAGGACACAGAGAAACTGGTAGCTTTCTTGTTTGGAATTACT  480
395  TATCCCCTTCCCTGGGAGAAGGACACAGAGAAACTGGTAGCTTTCTTGTTTGGAATTACT  454
     ------------------------------------------------------------

370  TCTCACATGGCGGCAGATGTCAGCTGGCATAGTCTGGGCCTTGAACAAGGATTCCTTAGG  429
481  TCTCACATGGCGGCAGATGTCAGCTGGCATAGTCTGGGCCTTGAACAAGGATTCCTTAGG  540
541  TCTCACATGGCGGCAGATGTCAGCTGGCATAGTCTGGGCCTTGAACAAGGATTCCTTAGG  514
     ------------------------------------------------------------

430  ACCATGGGAGCTATTGATTTTCACGGCTCCTATTCAGAGGCTCATTCGGCTGGTGATTTT  489
541  ACCATGGGAGCTATTGATTTTCACGGCTCCTATTCAGAGGCTCATTCGGCTGGTGATTTT  600
515  ACCATGGGAGCTATTGATTTTCACGGCTCCTATTCAGAGGCTCATTCGGCTGGTGATTTT  574
     ------------------------------------------------------------
```

```
490  GGAGGAGATGTGTTGAGCCAGTTTGAATTTAATTTTAATTACCTTGCACGACGCTGGTAT  549
601  GGAGGAGATGTGTTGAGCCAGTTTGAATTTAATTTTAATTACCTTGCACGACGCTGGTAT  660
575  GGAGGAGATGTGTTGAGCCAGTTTGAATTTAATTTTAATTACCTTGCACGACGCTGGTAT  634
     ------------------------------------------------------------

550  GTGCCAGTCAAAGATCTACTGGGAATTTATGAGAAACTGTATGGTCGAAAAGTCATCACC  609
661  GTGCCAGTCAAAGATCTACTGGGAATTTATGAGAAACTGTATGGTCGAAAAGTCATCACC  720
635  GTGCCAGTCAAAGATCTACTGGGAATTTATGAGAAACTGTATGGTCGAAAAGTCATCACC  694
     ------------------------------------------------------------

610  GAAAATGTAATCGTTGATTGTTCACATATCCAGTTCTTAGAAATGTATGGTGAGATGCTA  669
721  GAAAATGTAATCGTTGATTGTTCACATATCCAGTTCTTAGAAATGTATGGTGAGATGCTA  780
695  GAAAATGTAATCGTTGATTGTTCACATATCCAGTTCTTAGAAATGTATGGTGAGATGCTA  754
     ------------------------------------------------------------

670  GCTGTTTCCAAGTTATATCCCACTTACTCTACAAAGTCCCCGTTTTTGGTGGAACAATTC  729
781  GCTGTTTCCAAGTTATATCCCACTTACTCTACAAAGTCCCCGTTTTTGGTGGAACAATTC  840
755  GCTGTTTCCAAGTTATATCCCACTTACTCTACAAAGTCCCCGTTTTTGGTGGAACAATTC  814
     ------------------------------------------------------------

730  CAAGAGTATTTTCTTGGAGGACTGGATGATATGGCATTTTGGTCCACTAATATTTACCAT  789
841  CAAGAGTATTTTCTTGGAGGACTGGATGATATGGCATTTTGGTCCACTAATATTTACCAT  900
815  CAAGAGTATTTTCTTGGAGGACTGGATGATATGGCATTTTGGTCCACTAATATTTACCAT  874
     --------------------------------------------------GGGCTGTAAC  10

790  CTAACAATCTTCATGTTGGAGAATGGGACCAGTGACTGCAACCTGCCTGAGAACCCTCTG  849
901  CTAACAAGCTTCATGTTGGAGAATGGGACCAGTGACTGCAACCTGCCTGAGAACCCTCTG  960
875  CTAACAAGCTTCATGTTGGAGAATGGGACCAGTGACTGCAACCTGCCTGAGAACCCTCTG  934
11   TCTGCCATCCCTCAGCATAATTTGGGGGTATGATTTCACTATCCTAATTGCCTGTCCTAA  70

850  TTCATTGCATGTGGCGGCCAGCAAAACCACACCCAGGGCTCAAAAATGCAGAAAAAATGAT  909
961  TTCATTGCATGTGGCGGCCAGCAAAACCACACCCAGGGCTCAAAAATGCAGAAAAAATGAT  1020
935  TTCATTGCATGTGGCGGCCAGCAAAACCACACCCAGGGCTCAAAAATGCAGAAAAAATGAT  994
71   GTGATCTTACTTGCTGATAGGACCTAATGTTTATTTTATTGTTTAGCACTTCTAAAAAC  130

910  TTTCACAGAAATTTGACTACATCCCTAACTGAAAGTGTTGACAGGAATATAAACTATACT  969
1021 TTTCACAGAAATTTGACTACATCCCTAACTGAAAGTGTTGACAGGAATATAAACTATACT  1080
995  TTTCACAGAAATTTGACTACATCCCTAACTGAAAGTGTTGACAGGAATATAAACTATACT  1054
131  TCATTTCCTTTACACAAGTCCAATACTTTGGACAGGAAACAGTAGCTTTGTTGATTATGC  180

970  GAAAGAGGAGTGTTCTTTAGTGTAAATTCCTGGACCCCGGATTCCATGTCCTTTATCTAC  1029
1081 GAAAGAGGAGTGTTCTTTAGTGTAAATTCCTGGACCCCGGATTCCATGTCCTTTATCTAC  1140
1055 GAAAGAGGAGTGTTCTTTAGTGTAAATTCCTGGACCCCGGATTCCATGTCCTTTATCTAC  1114
181  TACGTGTCTTTACTGTCTATAATGATTCTTTTATTTCAGGATTCCATGTCCTTTATCTAC  240

1030 AAGGCTTTGGAAAGGAACATAAGGACAATGTTCATAGGTGGCTCTCAGTTGTCACAAAAG  1089
1141 AAGGCTTTGGAAAGGAACATAAGGACAATGTTCATAGGTGGCTCTCAGTTGTCACAAAAG  1200
1115 AAGGCTTTGGAAAGGAACATAAGGACAATGTTCATAGGTGGCTCTCAGTTGTCACAAAAG  1174
241  AAGGCTTTGGAAAGGAACATAAGGACAATGTTCATAGGTGGCTCTCAGTTGTCACAAAAG  300

1090 CACGTCTCCAGCCCCTTAGCATCTTACTTCTTGTCATTTCCTTATGCGAGGCTTGGCTGG  1149
1201 CACGTCTCCAGCCCCTTAGCATCTTACTTCTTGTCATTTCCTTATGCGAGGCTTGGCTGG  1260
1175 CACGTCTCCAGCCCCTTAGCATCTTACTTCTTGTCATTTCCTTATGCGAGGCTTGGCTGG  1234
301  CACGTCTCCAGCCCCTTAGCATCTTACTTCTTGTCATTTCCTTATGCGAGGCTTGGCTGG  360
```

```
1150 GCAATGACCTCAGCTGACCTCAACCAGGATGGGCACGGTGACCTCGTGGTGGGCGCACCA 1209
1261 GCAATGACCTCAGCTGACCTCAACCAGGATGGGCACGGTGACCTCGTGGTGGGCGCACCA 1320
1235 GCAATGACCTCAGCTGACCTCAACCAGGATGGGCACGGTGACCTCGTGGTGGGCGCACCA 1294
361  GCAATGACCTCAGCTGACCTCAACCAGGATGGGCACGGTGACCTCGTGGTGGGCGCACCA 420

1210 GGCTACAGCCGCCCCGGCCACATCCACATCGGGCGCGTGTACCTCATCTACGGCAATGAC 1269
1321 GGCTACAGCCGCCCCGGCCACATCCACATCGGGCGCGTGTACCTCATCTACGGCAATGAC 1380
1295 GGCTACAGCCGCCCCGGCCACATCCACATCGGGCGCGTGTACCTCATCTACGGCAATGAC 1354
421  GGCTACAGCCGCCCCGGCCACATCCACATCGGGCGCGTGTACCTCATCTACGGCAATGAC 480

1270 CTGGGCCTGCCACCTGTTGACCTGGACCTGGACAAGGAGGCCCACAGGATCCTTGAAGGC 1329
1381 CTGGGCCTGCCACCTGTTGACCTGGACCTGGACAAGGAGGCCCACAGGATCCTTGAAGGC 1440
1355 CTGGGCCTGCCACCTGTTGACCTGGACCTGGACAAGGAGGCCCACAGGATCCTTGAAGGC 1414
481  CTGGGCCTGCCACCTGTTGACCTGGACCTGGACAAGGAGGCCCACAGGATCCTTGAAGGC 540

1330 TTCCAGCCCTCAGGTCGGTTTGGCTCGGCCTTGGCTGTGTTGGACTTTAACGTGGACGGC 1389
1441 TTCCAGCCCTCAGGTCGGTTTGGCTCGGCCTTGGCTGTGTTGGACTTTAACGTGGACGGC 1500
1415 TTCCAGCCCTCAGGTCGGTTTGGCTCGGCCTTGGCTGTGTTGGACTTTAACGTGGACGGC 1474
541  TTCCAGCCCTCAGGTCGGTTTGGCTCGGCCTTGGCTGTGTTGGACTTTAACGTGGACGGC 600

1390 GTGCCTGACCTGGCCGTGGGAGCTCCCTCGGTGGGCTCCGAGCAGCTCACCTACAAAGGT 1449
1501 GTGCCTGACCTGGCCGTGGGAGCTCCCTCGGTGGGCTCCGAGCAGCTCACCTACAAAGGT 1560
1475 GTGCCTGACCTGGCCGTGGGAGCTCCCTCGGTGGGCTCCGAGCAGCTCACCTACAAAGGT 1534
601  GTGCCTGACCTGGCCGTGGGAGCTCCCTCGGTGGGCTCCGAGCAGCTCACCTACAAAGGT 660

1450 GCCGTGTATGTCTACTTTGGTTCCAAACAAGGAGGAATGTCTTCTTCCCCTAACATCACC 1509
1561 GCCGTGTATGTCTACTTTGGTTCCAAACAAGGAGGAATGTCTTCTTCCCCTAACATCACC 1620
1535 GCCGTGTATGTCTACTTTGGTTCCAAACAAGGAGGAATGTCTTCTTCCCCTAACATCACC 1594
661  GCCGTGTATGTCTACTTTGGTTCCAAACAAGGAGGAATGTCTTCTTCCCCTAACATCACC 720

1510 ATTTCTTGCCAGGACATCTACTGTAACTTGGGCTGGACTCTCTTGGCTGCAGATGTGAAT 1569
1621 ATTTCTTGCCAGGACATCTACTGTAACTTGGGCTGGACTCTCTTGGCTGCAGATGTGAAT 1680
1595 ATTTCTTGCCAGGACATCTACTGTAACTTGGGCTGGACTCTCTTGGCTGCAGATGTGAAT 1654
721  ATTTCTTGCCAGGACATCTACTGTAACTTGGGCTGGACTCTCTTGGCTGCAGATGTGAAT 780

1570 GGAGACAGTGAACCCGATCTGGTCATCGGCTCCCCTTTTGCACCAGGTGGAGGGAAGCAG 1629
1681 GGAGACAGTGAACCCGATCTGGTCATCGGCTCCCCTTTTGCACCAGGTGGAGGGAAGCAG 1740
1655 GGAGACAGTGAACCCGATCTGGTCATCGGCTCCCCTTTTGCACCAGGTGGAGGGAAGCAG 1714
781  GGAGACAGTGAACCCGATCTGGTCATCGGCTCCCCTTTTGCACCAGGTGGAGGGAAGCAG 840

1630 AAGGGAATTGTGGCTGCGTTTTATTCTGGCCCCAGCCTGAGCGACAAAGAAAAACTGAAC 1689
1741 AAGGGAATTGTGGCTGCGTTTTATTCTGGCCCCAGCCTGAGCGACAAAGAAAAACTGAAC 1800
1715 AAGGGAATTGTGGCTGCGTTTTATTCTGGCCCCAGCCTGAGCGACAAAGAAAAACTGAAC 1774
841  AAGGGAATTGTGGCTGCGTTTTATTCTGGCCCCAGCCTGAGCGACAAAGAAAAACTGAAC 900

1690 GTGGAGGCAGCCAACTGGACGGTGAGAGGCGAGGAAGACTTCTCCTGGTTTGGATATTCC 1749
1801 GTGGAGGCAGCCAACTGGACGGTGAGAGGCGAGGAAGACTTCTCCTGGTTTGGATATTCC 1860
1775 GTGGAGGCAGCCAACTGGACGGTGAGAGGCGAGGAAGACTTCTCCTGGTTTGGATATTCC 1834
901  GTGGAGGCAGCCAACTGGACGGTGAGAGGCGAGGAAGACTTCTCCTGGTTTGGATATTCC 960

1750 CTTCACGGTGTCACTGTGGACAACAGAACCTTGCTGTTGGTTGGGAGCCCGACCTGGAAG 1809
1861 CTTCACGGTGTCACTGTGGACAACAGAACCTTGCTGTTGGTTGGGAGCCCGACCTGGAAG 1920
1835 CTTCACGGTGTCACTGTGGACAACAGAACCTTGCTGTTGGTTGGGAGCCCGACCTGGAAG 1894
961  CTTCACGGTGTCACTGTGGACAACAGAACCTTGCTGTTGGTTGGGAGCCCGACCTGGAAG 1020
```

1810 AATGCCAGCAGGCTGGGCCATTTGTTACACATCCGAGATGAGAAAAAGAGCCTTGGGAGG 1869
1921 AATGCCAGCAGGCTGGGCCATTTGTTACACATCCGAGATGAGAAAAAGAGCCTTGGGAGG 1980
1895 AATGCCAGCAGGCTGGGCCATTTGTTACACATCCGAGATGAGAAAAAGAGCCTTGGGAGG 1954
1021 AATGCCAGCAGGCTGGGCCATTTGTTACACATCCGAGATGAGAAAAAGAGCCTTGGGAGG 1080

1870 GTGTATGGCTACTTCCCACCAAACGGCCAAAGCTGGTTTACCATTTCTGGAGACAAGGCA 1929
1981 GTGTATGGCTACTTCCCACCAAACGGCCAAAGCTGGTTTACCATTTCTGGAGACAAGGCA 2040
1955 GTGTATGGCTACTTCCCACCAAACGGCCAAAGCTGGTTTACCATTTCTGGAGACAAGGCA 2014
1081 GTGTATGGCTACTTCCCACCAAACGGCCAAAGCTGGTTTACCATTTCTGGAGACAAGGCA 1140

1930 ATGGGGAAACTGGGTACTTCCCTTTCCAGTGGCCACGTACTGATGAATGGGACTCTGAAA 1989
2041 ATGGGGAAACTGGGTACTTCCCTTTCCAGTGGCCACGTACTGATGAATGGGACTCTGAAA 2100
2015 ATGGGGAAACTGGGTACTTCCCTTTCCAGTGGCCACGTACTGATGAATGGGACTCTGAAA 2074
1141 ATGGGGAAACTGGGTACTTCCCTTTCCAGTGGCCACGTACTGATGAATGGGACTCTGAAA 1200

1990 CAAGTGCTGCTGGTTGGAGCCCCTACGTACGATGACGTGTCTAAGGTGGCATTCCTGACC 2049
2101 CAAGTGCTGCTGGTTGGAGCCCCTACGTACGATGACGTGTCTAAGGTGGCATTCCTGACC 2160
2075 CAAGTGCTGCTGGTTGGAGCCCCTACGTACGATGACGTGTCTAAGGTGGCATTCCTGACC 2134
1201 CAAGTGCTGCTGGTTGGAGCCCCTACGTACGATGACGTGTCTAAGGTGGCATTCCTGACC 1260

2050 GTGACCCTACACCAAGGCGGAGCCACTCGCATGTACGCACTCATATCTGACGCGCAGCCT 2109
2161 GTGACCCTACACCAAGGCGGAGCCACTCGCATGTACGCACTCATATCTGACGCGCAGCCT 2220
2135 GTGACCCTACACCAAGGCGGAGCCACTCGCATGTACGCACTCATATCTGACGCGCAGCCT 2194
1261 GTGACCCTACACCAAGGCGGAGCCACTCGCATGTACGCACTCATATCTGACGCGCAGCCT 1320

2110 CTGCTGCTCAGCACCTTCAGCGGAGACCGCCGCTTCTCCCGATTTGGTGGCGTTCTGCAC 2169
2221 CTGCTGCTCAGCACCTTCAGCGGAGACCGCCGCTTCTCCCGATTTGGTGGCGTTCTGCAC 2280
2195 CTGCTGCTCAGCACCTTCAGCGGAGACCGCCGCTTCTCCCGATTTGGTGGCGTTCTGCAC 2254
1321 CTGCTGCTCAGCACCTTCAGCGGAGACCGCCGCTTCTCCCGATTTGGTGGCGTTCTGCAC 1380

2170 TTGAGTGACCTGGATGATGATGGCTTAGATGAAATCATCATGGCAGCCCCCCTGAGGATA 2229
2281 TTGAGTGACCTGGATGATGATGGCTTAGATGAAATCATCATGGCAGCCCCCCTGAGGATA 2340
2255 TTGAGTGACCTGGATGATGATGGCTTAGATGAAATCATCATGGCAGCCCCCCTGAGGATA 2314
1381 TTGAGTGACCTGGATGATGATGGCTTAGATGAAATCATCATGGCAGCCCCCCTGAGGATA 1440

2230 GCAGATGTAACCTCTGGACTGATTGGGGGAGAAGACGGCCGAGTATATGTATATAATGGC 2289
2341 GCAGATGTAACCTCTGGACTGATTGGGGGAGAAGACGGCCGAGTATATGTATATAATGGC 2400
2315 GCAGATGTAACCTCTGGACTGATTGGGGGAGAAGACGGCCGAGTATATGTATATAATGGC 2374
1441 GCAGATGTAACCTCTGGACTGATTGGGGGAGAAGACGGCCGAGTATATGTATATAATGGC 1500

2290 AAAGAGACCACCCTTGGTGACATGACTGGCAAATGCAAATCATGGATAACTCCATGTCCA 2349
2401 AAAGAGACCACCCTTGGTGACATGACTGGCAAATGCAAATCATGGATAACTCCATGTCCA 2460
2375 AAAGAGACCACCCTTGGTGACATGACTGGCAAATGCAAATCATGGATAACTCCATGTCCA 2434
1501 AAAGAGACCACCCTTGGTGACATGACTGGCAAATGCAAATCATGGATAACTCCATGTCCA 1560

2350 GAAGAAAAGGCCCAATATGTATTGATTTCTCCTGAAGCCAGCTCAAGGTTTGGGAGCTCC 2409
2461 GAAGAAAAGGCCCAATATGTATTGATTTCTCCTGAAGCCAGCTCAAGGTTTGGGAGCTCC 2520
2435 GAAGAAAAGGTAAGTGAAAAAAAAAAAAAAAAAAAAAA---------------------- 2472
1561 GAAGAAAAGGCCCAATATGTATTGATTTCTCCTGAAGCCAGCTCAAGGTTTGGGAGCTCC 1620

2410 CTCATCACCGTGAGGTCCAAGGCAAAGAACCAAGTCGTCATTGCTGCTGGAAGGAGTTCT 2469
2521 CTCATCACCGTGAGGTCCAAGGCAAAGAACCAAGTCGTCATTGCTGCTGGAAGGAGTTCT 2580

1621 CTCATCACCGTGAGGTCCAAGGCAAAGAACCAAGTCGTCATTGCTGCTGGAAGGAGTTCT 1680

8.5

17118

```
2470 TTGGGAGCCCGACTCTCCGGGGCACTTCACGTCTATAGCCTTGGCTCAGATTGAAGATTT 2529
2581 TTGGGAGCCCGACTCTCCGGGGCACTTCACGTCTATAGCCTTGGCTCAGATTGAAGATTT 2640
     ------------------------------------------------------------
1681 TTGGGAGCCCGACTCTCCGGGGCACTTCACGTCTATAGCCTTGGCTCAGATTGAAGATTT 1740

2530 CACTGCATTTCCCCACTCTGCCCACCTCTCTCATGCTGAATCACATCCATGGTGAGCATT 2589
2641 CACTGCATTTCCCCACTCTGCCCACCTCTCTCATGCTGAATCACATCCATGGTGAGCATT 2700
     ------------------------------------------------------------
1741 CACTGCATTTCCCCACTCTGCCCACCTCTCTCATGCTGAATCACATCCATGGTGAGCATT 1800

2590 TTGATGGACAAAGTGGCACATCCAGTGGAGCGGTGGTAGATCCTGATAGACATGGGGCTC 2649
2701 TTGATGGACAAAGTGGCACATCCAGTGGAGCGGTGGTAGATCCTGATAGACATGGGGCTC 2760
     ------------------------------------------------------------
1801 TTGATGGACAAAGTGGCACATCCAGTGGAGCGGTGGTAGATCCTGATAGACATGGGGCTC 1860

2650 CTGGGAGTAGAGAGACACACTAACAGCCACACCCTCTGGAAATCTGATACAGTAAATATA 2709
2761 CTGGGAGTAGAGAGACACACTAACAGCCACACCCTCTGGAAATCTGATACAGTAAATATA 2820
     ------------------------------------------------------------
1861 CTGGGAGTAGAGAGACACACTAACAGCCACACCCTCTGGAAATCTGATACAGTAAATATA 1920

2710 TGACTGCACCAGAAATATGTGAAATAGCAGACATTCTGCTTACTCATGTCTCCTTCCACA 2769
2821 TGACTGCACCAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2880
     --------------------------------                             1952
1921 TGACTGCACCAGAAAAAAAAAAAAAAAAAAAA

2770 GTTTACTTCCTCGCTCCCTTTGCATCTAAACCTTTCTTCTTTCCCAACTTATTGCCTGTA 2829
2881 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA------------------------- 2915
     ------------------------------------------------------------
     ------------------------------------------------------------

2830 GTCAGACCTGCTGTACAACCTATTTCCTCTTCCTCTTGAATGTCTTTCCAGTGGCTGGAA 2889
     ------------------------------------------------------------
     ------------------------------------------------------------
     ------------------------------------------------------------

2890 AGGTCCCTCTGTGGTTATCTGTTAGAACAGTCTCTGTACACAATTCCTCCTAAAAACATC 2949
     ------------------------------------------------------------
     ------------------------------------------------------------
     ------------------------------------------------------------

2950 CTTTTTTAAAAAAAGAATTGTTCAGCCATAAAGAAAGAACAAGATCATGCCCTTTGCAGG 3009
     ------------------------------------------------------------
     ------------------------------------------------------------
     ------------------------------------------------------------

3010 GACATGGATGGAGCTGGAGGCCATTATCCTTCATAAACTATTGCAGGAACAGAAAACCAA 3069
     ------------------------------------------------------------
     ------------------------------------------------------------
     ------------------------------------------------------------

3070 ACACTCCATATTCTCACTTGTAAGTGGGAGCTAAGTGAGAACACGTGGACACATAGAGGG 3129
     ------------------------------------------------------------
     ------------------------------------------------------------
     ------------------------------------------------------------
```

0.6

```
3130 AAACAACACACACTGGGGCCTATGAGAGGGCGGAAGGTGGGAGGAGGGAGAGATCAGGAA 3189
     ------------------------------------------------------------
     ------------------------------------------------------------
     ------------------------------------------------------------

3190 AAATAACTAATGGATACTTAGGGTGATGAAATAATCTGTGTAACAAACCCCCATGACACA 3249
     ------------------------------------------------------------
     ------------------------------------------------------------
     ------------------------------------------------------------

3250 CCTTTATGTATGTAACAAACCAGCACTTCCTGCGCATGTACCCCTGAACTTAAAAGTTAA 3309
     ------------------------------------------------------------
     ------------------------------------------------------------
     ------------------------------------------------------------

3310 AAAAAAGTTGAACTTAAAAATAACAGATTGGCCCATGCCAATCAAAGTATAATAGAAAGC 3369
     ------------------------------------------------------------
     ------------------------------------------------------------
     ------------------------------------------------------------

3370 ATAGTATAC 3378
     ---------
     ---------
     ---------
```

PCT NO : GB99 / 04399

Form 23/77 : 24/12/99

Agent : Newburn Ellis